

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:37:30 ; Search time 75.6667 Seconds

(without alignments)
636.206 Million cell updates/sec

Title: US-10-705-716A-4

Perfect score: 1 MCGGSRADAIERYRYESWT.....VTDISIQMDRSRRITKNCVN 145

Sequence:

BLOSUM62

Scoring table: Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Database:

Published Applications AA:*

1: /cgn2_6/ptodata/1/pubppa/US07_PUBCOMB.pep.*
2: /cgn2_6/ptodata/1/pubppa/PTCT_NEW_PUB.pep.*
3: /cgn2_6/ptodata/1/pubppa/US06_NEW_PUB.pep.*
4: /cgn2_6/ptodata/1/pubppa/US06_PUBCOMB.pep.*
5: /cgn2_6/ptodata/1/pubppa/US07_NEW_PUB.pep.*
6: /cgn2_6/ptodata/1/pubppa/PTCT_PUBCOMB.pep.*
7: /cgn2_6/ptodata/1/pubppa/US08_NEW_PUB.pep.*
8: /cgn2_6/ptodata/1/pubppa/US08_PUBCOMB.pep.*
9: /cgn2_6/ptodata/1/pubppa/US09_PUBCOMB.pep.*
10: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
11: /cgn2_6/ptodata/1/pubppa/US09C_PUBCOMB.pep.*
12: /cgn2_6/ptodata/1/pubppa/US09_NEW_PUB.pep.*
13: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
14: /cgn2_6/ptodata/1/pubppa/US10_PUBCOMB.pep.*
15: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
16: /cgn2_6/ptodata/1/pubppa/US10C_PUBCOMB.pep.*
17: /cgn2_6/ptodata/1/pubppa/US10_NEW_PUB.pep.*
18: /cgn2_6/ptodata/1/pubppa/US11_NEW_PUB.pep.*
19: /cgn2_6/ptodata/1/pubppa/US60_NEW_PUB.pep.*
20: /cgn2_6/ptodata/1/pubppa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	100.0	145	US-10-293-239-17	Sequence 17, Appl
2	778	100.0	145	US-10-177-390-30	Sequence 30, Appl
3	750.5	96.5	180	US-10-293-239-19	Sequence 19, Appl
4	569.5	73.2	149	US-10-293-239-21	Sequence 21, Appl
5	309	38.4	73	US-10-293-239-20	Sequence 20, Appl
6	299	38.4	54	US-10-293-239-18	Sequence 18, Appl
7	293	37.7	80	US-10-293-239-22	Sequence 22, Appl
8	140	18.0	25	US-10-293-239-37	Sequence 37, Appl
9	99	12.7	18	US-10-293-239-35	Sequence 35, Appl
10	92.5	11.9	307	US-10-437-963-181279	Sequence 181279,
11	92.5	11.9	1001	US-10-415-147-3	Sequence 3, Appl1
12	91	11.7	670	US-10-156-761-14107	Sequence 14107, A
13	90	11.6	219	US-10-156-761-13447	Sequence 13447, A

14	88	11.3	298	US-09-910-600-26	Sequence 26, Appl
15	87.5	11.2	369	US-10-036-547-84	Sequence 84, Appl
16	87.5	11.2	369	US-10-188-832-149	Sequence 149, Appl
17	87.5	11.2	369	US-10-658-884-4	Sequence 4, Appl1
18	87.5	11.2	383	US-10-029-386-32058	Sequence 32058, A
19	87	11.2	136	US-10-424-599-179100	Sequence 179100,
20	86	11.1	182	US-10-767-701-37779	Sequence 37779, A
21	86	11.1	342	US-10-162-944-914	Sequence 914, App
22	86	11.1	668	US-10-437-963-134726	Sequence 134726,
23	85.5	11.0	466	US-10-425-114-71015	Sequence 71015, A
24	85	10.9	486	US-10-282-122A-62114	Sequence 62114, A
25	84	10.8	514	US-09-712-363-154	Sequence 154, App
26	83.5	10.7	332	US-10-437-963-168390	Sequence 168390,
27	83.5	10.7	558	US-10-156-761-12110	Sequence 12110, A
28	83	10.7	200	US-10-767-701-32315	Sequence 32315, A
29	83	10.7	373	US-10-437-963-147910	Sequence 147910,
30	83	10.7	2483	US-10-741-600-1170	Sequence 1170, Ap
31	82.5	10.6	497	US-10-437-963-157852	Sequence 157852,
32	82.5	10.6	850	US-10-424-599-242653	Sequence 242653,
33	82	10.5	216	US-10-767-701-57343	Sequence 57343, A
34	82	10.5	657	US-10-437-963-163549	Sequence 163549,
35	82	10.5	795	US-10-424-599-174901	Sequence 174901,
36	81.5	10.5	326	US-10-425-114-69350	Sequence 69350, A
37	81	10.4	147	US-10-767-701-46905	Sequence 46905, A
38	81	10.4	243	US-10-424-599-209325	Sequence 209325,
39	81	10.4	355	US-10-425-114-42733	Sequence 42733, A
40	81	10.4	2527	US-10-408-765A-2462	Sequence 2462, Ap
41	81	10.4	2715	US-10-408-765A-866	Sequence 866, App
42	80.5	10.3	213	US-10-424-599-190791	Sequence 190791,
43	80.5	10.3	337	US-10-437-963-107097	Sequence 107097,
44	80.5	10.3	342	US-09-269-390-4	Sequence 4, Appl1
45	80.5	10.3	342	US-10-317-832-118	Sequence 118, App

ALIGNMENTS

RESULT 1	US-10-293-239-17	Application US/10293239
Sequence 17,	Application No. (US20030119043A1)	
GENERAL INFORMATION:		
APPLICANT:	Tanner, Stephan	
APPLICANT:	de la Chapelle, Albert	
TITLE OF INVENTION:	BALC expression as a diagnostic marker for acute leukemia	
FILE REFERENCE:	22727/04101	
CURRENT APPLICATION NUMBER:	US/10/293,239	
CURRENT FILING DATE:	2002-11-12	
PRIOR APPLICATION NUMBER:	US 60/346,210	
PRIOR FILING DATE:	2001-11-09	
NUMBER OF SEQ ID NOS:	39	
SOFTWARE:	PatentIn version 3.1	
SEQ ID NO 17		
LENGTH:	145	
TYPE:	PRT	
ORGANISM:	Homo sapiens	
US-10-293-239-17		
Query Match	Score 778.0; DB 14; Length 145;	
Best Local Similarity	100.0%; Pred. No. 1e-66;	
Matches	145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 MCGGSRADAIERYRYESWTRETSTWLYTSDAPPSAAPPSPGAGLHSGMLDGL 60	
DB	1 MCGGSRADAIERYRYESWTRETSTWLYTSDAPPSAAPPSPGAGLHSGMLDGL 60	
QY	PSNGVRSRTAPGGIIPMPKKTNCETOCPPNOSLSSGPLQKQKGLQTTAKRKPMPAK 120	
DB	PSNGVRSRTAPGGIIPMPKKTNCETOCPPNOSLSSGPLQKQKGLQTTAKRKPMPAK 120	
QY	EVTINVTDSIQMDRSRRITKNCVN 145	
DB	EVTINVTDSIQMDRSRRITKNCVN 145	

RESULT 2

US-10-177-390-30
; Sequence 30, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerald
; APPLICANT: N.V. Antwerp Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear
; FILE REFERENCE: 021505wc/JH/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-30

Query Match 100.0%; Score 778; DB 14; Length 145;
Best Local Similarity 100.0%; Pred. No. 1e-66;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 60
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 60

QY 61 PSNGVPRSTAPGIPNPEKKTNCETOCNPNQSLSSGPLTQKONGLOTTEAKRDAKMPAK 120
DB 61 PSNGVPRSTAPGIPNPEKKTNCETOCNPNQSLSSGPLTQKONGLOTTEAKRDAKMPAK 120

QY 121 EYTIINTDSIQOMDRSRRTKNCVN 145
DB 121 EYTIINTDSIQOMDRSRRTKNCVN 145

RESULT 3

US-10-293-239-19
; Sequence 19, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-19

Query Match 96.5%; Score 750.5; DB 14; Length 180;
Best Local Similarity 80.6%; Pred. No. 6e-64;
Matches 145; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 53
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 60

QY 54 -----GMLEDDGIPNCGVPRSTAPGIPNPEKKTNCET 85
DB 61 KIKAPLDSVSDGLFSASKMAPLAVFSHGMLEDDGIPNCGVPRSTAPGIPNPEKKTNCET 120

QY 86 QCPNPOSLSGPLTQKONGLOTTEAKRDAKMPAKVNTINTDSIQOMDRSRRTKNCVN 145

DB

121 QCPNPOSLSGPLTQKONGLOTTEAKRDAKMPAKVNTINTDSIQOMDRSRRTKNCVN 180

RESULT 4

US-10-293-239-21
; Sequence 21, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-21

Query Match 73.2%; Score 569.5; DB 14; Length 149;
Best Local Similarity 75.7%; Pred. No. 1.1e-46;
Matches 109; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 53
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 60

QY 54 -----GMLEDDGIPNCGVPRSTAPGIPNPEKKTNCET 85
DB 61 KIKAPLDSVSDGLFSASKMAPLAVFSHGMLEDDGIPNCGVPRSTAPGIPNPEKKTNCET 120

QY 86 QCPNPOSLSGPLTQKONGLOTTE 109
DB 121 QCPNPOSLSGPLTQKONGLOTTE 144

RESULT 5

US-10-293-239-20
; Sequence 20, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-20

Query Match 39.7%; Score 309; DB 14; Length 73;
Best Local Similarity 96.6%; Pred. No. 4.6e-22;
Matches 56; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 58
DB 1 MCGGSRADAIERPRYESWTRTESTWLTYYTDSADAPPSAAPPDSGPEAGLHSGMLEDDGI 58

RESULT 6

```
US-10-293-239-18
; Sequence 18, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 18
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-18

Query Match
Best Local Similarity 100.0%; Score 299; DB 14; Length 54;
Pred. No. 2.9e-21; Mismatches 0; Indels 0; Gaps 0;
Matches 54; Conservative 0;

Qy
1 MCGGSRADAIEPRYESWTRETESTWLTYYTDSADAPPAAAPDSGPEAGGLHS 54
Db
1 MCGGSRADAIEPRYESWTRETESTWLTYYTDSADAPPAAAPDSGPEAGGLHS 54

RESULT 7
US-10-293-239-22
; Sequence 22, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-22

Query Match
Best Local Similarity 100.0%; Score 293; DB 14; Length 80;
Pred. No. 1.8e-20; Mismatches 0; Indels 0; Gaps 0;
Matches 53; Conservative 0;

Qy
1 MCGGSRADAIEPRYESWTRETESTWLTYYTDSADAPPAAAPDSGPEAGGLHS 53
Db
1 MCGGSRADAIEPRYESWTRETESTWLTYYTDSADAPPAAAPDSGPEAGGLHS 53

RESULT 8
US-10-293-239-37
; Sequence 37, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 37
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-37

Query Match
Best Local Similarity 100.0%; Score 293; DB 14; Length 54;
Pred. No. 2.9e-21; Mismatches 0; Indels 0; Gaps 0;
Matches 54; Conservative 0;

Qy
1 MCGGSRADAIEPRYESWTRETESTWLTYYTDSADAPPAAAPDSGPEAGGLHS 54
Db
1 MCGGSRADAIEPRYESWTRETESTWLTYYTDSADAPPAAAPDSGPEAGGLHS 54

RESULT 9
US-10-293-239-35
; Sequence 35, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 35
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-35

Query Match
Best Local Similarity 100.0%; Score 99; DB 14; Length 18;
Pred. No. 0.011; Mismatches 0; Indels 0; Gaps 0;
Matches 18; Conservative 0;

Qy
9 DAIEPRYESWTRETEST 26
Db
1 DAIEPRYESWTRETEST 18

RESULT 10
US-10-437-963-181279
; Sequence 181279, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 181279
; LENGTH: 307
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(307)
; OTHER INFORMATION: unsure at all Xaa locations
```

RESULT 12
US-10-156-761-14107
; Sequence 14107, Applicant US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, JUN
; APPLICANT: HORIKAWA, HIROSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI

```
Matches      20; Conservative      8; Mismatches      31; Indels      0; Gaps      0;
Qy          6 SRADATIEPRRYESWTRFTESTLWTYTDSDAPPASAAPDGSBEAGGLSHGMLDEGLPSNG 64
           :|||:::||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db          43 ARRGVGNQTITRRMWSKAEVLFAFTDLSPQAALAAARPEEGBGGEGGCGENIIPDTG 101
```

```
/ GENERAL INFORMATION:
/ APPLICANT: Longphre, Malinda
/ APPLICANT: Chang, Han
/ APPLICANT: Whitney, Gena
/ TITLE OF INVENTION: NOVEL SIGLECS AND USES THEREOF
/ FILE REFERENCE: D0003NP
/ CURRENT APPLICATION NUMBER: US/09/910,600
/ CURRENT FILING DATE: 2001-07-20
/ PRIOR APPLICATION NUMBER: 60/220,139
/ PRIOR FILING DATE: 2000-07-21
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 26
/ LENGTH: 298
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: L3cyto-Y641
US-09-910-600-26
```

```
Query Match          11.3%; Score 88; DB 10; Length 298;
Best Local Similarity 27.5%; Pred. No. 4.2;
Matches 33; Conservative 18; Mismatches 43; Indels 26; Gaps 7;
```

```
QY 3 CGGSRADAI-----EPRYESMTRETSTW-LTYTDSAPPSA-----AAPDS---GP 46
DB 178 CFKKRIEALPQIKYLSKRTIAMPLO--GWOATFGGDHPKSDLVPRGSPNSINVP 234
QY 47 EAAGLHSGMLEDLPSNGVPRSTAPGIPNPEKKTNCETQ-----CPNPOSTSSGPLETOR 101
DB 235 TAGPLAQKXNQKATPNS--PRTPLPGAPSPESKXNQKQYQLSPFPEPKSTQAPESQE 292
```

```
RESULT 15
US-10-036-542-84
/ Sequence 84, Application US/10036542
/ Publication No. US20030083481A1
/ GENERAL INFORMATION:
/ APPLICANT: Blise et al.
/ TITLE OF INVENTION: 25 Human Prostate and Prostate Cancer Associated Proteins
/ FILE REFERENCE: PA002PI
/ CURRENT APPLICATION NUMBER: US/10/036,542
/ CURRENT FILING DATE: 2002-01-07
/ PRIOR APPLICATION NUMBER: PCT/US00/19666
/ PRIOR FILING DATE: 2000-07-20
/ PRIOR APPLICATION NUMBER: 60/144,972
/ PRIOR FILING DATE: 1999-07-21
/ PRIOR APPLICATION NUMBER: 60/148,681
/ PRIOR FILING DATE: 1999-08-13
/ PRIOR APPLICATION NUMBER: 60/149,173
/ PRIOR FILING DATE: 1999-08-17
/ PRIOR APPLICATION NUMBER: 60/158,004
/ PRIOR FILING DATE: 1999-10-06
/ PRIOR APPLICATION NUMBER: 60/194,689
/ PRIOR FILING DATE: 2000-04-05
/ NUMBER OF SEQ ID NOS: 157
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 84
/ LENGTH: 369
/ TYPE: PRT
/ ORGANISM: Homo sapiens
US-10-036-542-84
```

```
Query Match          11.2%; Score 87.5; DB 14; Length 369;
Best Local Similarity 25.9%; Pred. No. 6.2;
Matches 28; Conservative 19; Mismatches 48; Indels 13; Gaps 4;
```

```
QY 33 SDAPPSAADSPGPEAGLHSGMLEDLPSNGVPRS-TAPGIPNPEKKTNCETQCPNPQ 91
DB 40 SSTSTSSSPSPSSSSSSSSCYPLPS--TPEVSADETPNPQ--SAQIACSSPS 95
QY 92 SLSSGFL-----TQKNGLOTTEAKRDAKMPAKVITINTDSIQ 131
```

```
DB 96 VVASLPLDQSDGSSSQKESSPSTLQVLPDSLSLPRSEIDEKVTDLVQ 143
```

```
Search completed: April 12, 2005, 16:03:42
Job time : 75.6667 secs
```

This Page Blank (uspto)

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension .rup) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

CC sequence and an oligonucleotide comprising a sequence complementary to a
 CC polynucleotide which comprises a 3'-end sequence, where the
 CC oligonucleotide comprises at least 15 nucleotides and the combination of
 CC the 5'-end sequence/3'-end sequence is selected from those defined in the
 CC specification. The primer sets can be used in antisense therapy and in
 CC gene therapy. The primers are useful for synthesizing polynucleotides,
 CC particularly full-length cDNAs. The primers are also useful for the
 CC detection and/or diagnosis of the abnormality of the proteins encoded by
 CC the full-length cDNAs. The primers allow obtaining of the full-length
 CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
 CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
 CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
 CC oligonucleotides, all of which are used in the exemplification of the
 CC present invention

CC Sequence 145 AA;

Query Match 100.0%; Score 778; DB 4; Length 145;
 Best Local Similarity 100.0%; Pred. No. 4e-70; Mismatches 0; Gaps 0;
 Matches 145; Conservative 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDL 60
 DB 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDL 60
 QY 61 PANGVPRSTAPGCIIPNPEKKTNCETQCPNPQSLSSGPTLTKONGLOTTEAKDKAKMPAK 120
 DB 61 PANGVPRSTAPGCIIPNPEKKTNCETQCPNPQSLSSGPTLTKONGLOTTEAKDKAKMPAK 120
 QY 121 EYTIWNTDSIQMDRSRRITTKNCVN 145
 DB 121 EYTIWNTDSIQMDRSRRITTKNCVN 145

RESULT 2

AA019498
 ID AA019498 standard; protein; 145 AA.

AC AA019498;

DT 20-DEC-2002 (first entry)

DE Hs1 protein variant.

XX Hs1; variant; cancer; tumour; ungine cluster; cytosolic; metastasis;
 KW E87; expressed sequence tag; colon cancer; stomach cancer; breast cancer;
 KW HS169395; HS127144; HS2; HS132793; HS3.

XX Unidentified.

XX DE10103694-A1.

XX 26-JAN-2001; 2001DE-01003694.

XX 26-JAN-2001; 2001DE-01003694.

XX 26-JAN-2001; 2001DE-01003694.

XX (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX Brecto, Kemmerer W;

XX WPI; 2002-644836/70.

XX N-PSDB; AAL50100.

PT Diagnosis and therapy of tumors, by determining expression rates of
 PT specific expressed sequence tags of the ungine cluster, and subsequently
 PT blocking their expression.

PS Claim 10; Page 5; 10pp; German.

CC The present invention relates to the use of expressed sequence tags
 (ESTs), or variants, of the ungine cluster HS169395 (HS1), HS127144

CC (HS2) and/or HS132793 (HS3) for diagnosis and therapy of tumors, in
 CC which their expression rates in tumour cells and/or lymph nodes are
 CC determined. The EST sequences are useful as prognostic markers of
 CC survival of cancer patients (high levels of EST-related mRNA are
 CC associated with a poor prognosis, specifically correlated with
 CC development of metastases); and for diagnosis and/or therapy of solid
 CC tumours, particularly of colon, stomach and breast. The present sequence
 CC is a variant of the Hs1 protein shown in the exemplification of the
 CC invention

XX Sequence 145 AA;

Query Match 100.0%; Score 778; DB 5; Length 145;
 Best Local Similarity 100.0%; Pred. No. 4e-70; Mismatches 0; Gaps 0;
 Matches 145; Conservative 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDL 60
 DB 1 MCGGSRADAIIPRYESWTRETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDL 60
 QY 61 PANGVPRSTAPGCIIPNPEKKTNCETQCPNPQSLSSGPTLTKONGLOTTEAKDKAKMPAK 120
 DB 61 PANGVPRSTAPGCIIPNPEKKTNCETQCPNPQSLSSGPTLTKONGLOTTEAKDKAKMPAK 120
 QY 121 EYTIWNTDSIQMDRSRRITTKNCVN 145
 DB 121 EYTIWNTDSIQMDRSRRITTKNCVN 145

RESULT 3

ABR58646
 ID ABR58646 standard; protein; 145 AA.

AC ABR58646;

DT 09-JUL-2003 (first entry)

DE Human cancer related protein SEQ ID NO:303.

XX Human; cancer; diagnosis; screening; modulator; leukaemia; leaemia;
 KW heart disease; atherosclerosis; endometriosia.

XX Homo sapiens.

XX WO2003025138-A2.

XX 27-MAR-2003.

XX 17-SEP-2002; 2002WO-US029560.

XX 17-SEP-2001; 2001US-0323469P.

XX 20-SEP-2001; 2001US-0323887P.

XX 13-NOV-2001; 2001US-0350665P.

XX 08-FEB-2002; 2002US-0355145P.

XX 08-FEB-2002; 2002US-0355257P.

XX 12-APR-2002; 2002US-0372246P.

XX (EOSB-) EOS BIOTECHNOLOGY INC.

XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE,
 PI Zlocnik A;

XX WPI; 2003-354600/33.

XX N-PSDB; ACC72796.

PT New genes that are up-regulated or down-regulated in cancers, useful as
 PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 PT therapeutic targets for screening drugs for treating these diseases.
 PS Claim 12; Page 753; 767pp; English.
 CC The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-

1.131 de c lank

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:15:13 ; Search time 103.667 Seconds
(without alignment)

540.967 Million cell updates/sec

Title: US-10-705-716a-4

Perfect score: 778
Sequence: 1 MCGGSRRAAIERPRVYEST.....VDSIQMDRSRRITKNCVN 145

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_16Dec04.*

1: geneseqp1980s.*
2: geneseqp1980s.*
3: geneseqp2000s.*
4: geneseqp2001s.*
5: geneseqp2002s.*
6: geneseqp2003as.*
7: geneseqp2003bs.*
8: geneseqp2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	778	100.0	145	4	AA895018 Aab95018 Human pro
2	778	100.0	145	5	AA019498 Aa019498 HSI prote
3	778	100.0	145	6	ABR58646 Abt58646 Human can
4	778	100.0	145	7	AD031800 Adc31800 Human nov
5	778	100.0	145	7	ADMA6959 Adm46959 Brain and
6	778	100.0	145	8	AD048475 Ado48475 Human PTH
7	750.5	96.5	180	7	ADMA6961 Adm46961 Brain and
8	649	83.4	145	8	AD048479 Ado48479 Mouse PTH
9	645	82.9	145	8	AD048473 Ado48473 Rat PTH I
10	569.5	73.2	149	7	ADMA6963 Adm46963 Brain and
11	405.5	52.1	92	5	AD081902 Adq81902 Human dlo
12	309	39.7	73	7	ADMA6962 Adm46962 Brain and
13	299	38.4	54	7	ADMA6960 Adm46960 Brain and
14	299	38.4	54	8	AD048477 Ado48477 Human PTH
15	293	37.7	80	7	ADMA6964 Adm46964 Brain and
16	278	35.7	54	8	AD048481 Ado48481 Mouse PTH
17	140	18.0	25	7	ADMA6979 Adm46979 Brain and
18	99	12.7	18	7	ADMA6977 Adm46977 Brain and
19	92.5	11.9	1001	5	AAU89903 Aau89903 Rat inos1
20	91.5	11.8	210	3	AAU85390 Aag15390 Arabidops
21	91.5	11.8	222	3	AAU85389 Aag15389 Arabidops
22	91.5	11.8	612	8	ADRB6172 Adrb6172 Aspergill
23	89	11.4	16	8	AD048482 Ado48482 PTH respo
24	89	11.4	321	4	AB871629 Abb71629 Drosophi1
25	88	11.3	298	5	AAU87066 Aau87066 Sg1lec-BM

26	87.5	11.2	369	3	AAU71485 Aay71485 Human MAG
27	87.5	11.2	369	4	AA800297 Aab80297 Human pro
28	87.5	11.2	369	6	ABR48215 Abt48215 Human bla
29	87.5	11.2	369	6	ABU56516 Abu56516 Lung canc
30	87.5	11.2	383	8	ABO58424 Abo58424 Human gen
31	86.5	11.1	1192	4	ABG02038 Abg02038 Novel hum
32	86	11.1	342	8	ADM87821 Adm87821 Human EST
33	86	11.1	508	4	ABG09910 Abg09910 Novel hum
34	86	11.1	586	5	AB804711 Abb04711 Human PPI
35	86	11.1	991	4	AA883195 Aab83195 Human Rec
36	86	11.1	991	8	ADO55153 Ado55153 Protein #
37	85.5	11.0	211	8	ABM81408 Abm81408 Tumour-as
38	85.5	11.0	236	8	AD510476 Ad510476 Human the
39	85	10.9	486	6	ABU34190 Abu34190 Protein e
40	84.5	10.9	1111	5	AA017108 Aa017108 Murine G1
41	84	10.8	514	4	AA681103 Aag81103 Mycobacte
42	83.5	10.7	684	4	AB869330 Abb69330 Drosophi1
43	83.5	10.7	899	4	AB865489 Abb65489 Drosophi1
44	83.5	10.7	899	4	AB865488 Abb65488 Drosophi1
45	83	10.7	659	4	AB865280 Abb65280 Drosophi1

ALIGNMENTS

RESULT-1
ID AAB95018 standard; protein; 145 AA.
XX
AC AAB95018;
XX
DT 26-UTN-2001 (first entry)

DE Human protein sequence SEQ ID NO:16726.

KM Human; primer; detection; diagnosis; antisense therapy; gene therapy.

OS Homo sapiens.

PN EP1074617-A2.

PD 07-FEB-2001.

XX 28-UTN-2000; 2000EP-00116126.

XX 29-JUL-1999; 99JP-00248036.

PR 27-AUG-1999; 99JP-00300253.

PR 11-JAN-2000; 2000JP-00118776.

PR 02-MAY-2000; 2000JP-00183767.

PR 09-UTN-2000; 2000JP-00241899.

XX (HELI-) HELIX RES INST.

PI Ota T, Isogai T, Nishikawa T, Hayashi K, Saito K, Yamamoto J;

PI Ishii S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;

XX WPI; 2001-318749/34.

PS Claim 8; SEQ ID NO 16726; 2537pp + Sequence Listing; English.

XX The present invention describes primer sets for synthesizing 5602 full-length cDNAs defined in the specification, where a primer set comprises:

CC (a) an oligo-dT primer and an oligonucleotide complementary to the complementary strand of a polynucleotide which comprises one of the 5602 complementary sequences defined in the specification, where the CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination of an oligonucleotide comprising a sequence complementary to the complementary strand of a polynucleotide which comprises a 5'-end

CC sequence and an oligonucleotide comprising a sequence complementary to a
CC polynucleotide which comprises a 3'-end nucleotide, where the
CC oligonucleotide comprises at least 15 nucleotides and the combination of
CC the 5'-end sequence/3'-end sequence is selected from those defined in the
CC specification. The primer sets can be used in antisense therapy and in
CC gene therapy. The primers are useful for synthesizing polynucleotides,
CC particularly full-length cDNAs. The primers are also useful for the
CC detection and/or diagnosis of the abnormality of the proteins encoded by
CC the full-length cDNAs. The primers allow obtaining of the full-length
CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
CC oligonucleotides, all of which are used in the exemplification of the
CC present invention
CC
XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 778; DB 4; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCGGSRADAIIPRYESWTRETSTWLTYYTDSADPPSAAPDSGPEAGLHSGMLEDGL 60
DB 1 MCGGSRADAIIPRYESWTRETSTWLTYYTDSADPPSAAPDSGPEAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGGINPEKKTNCETCCPNPQSLSSGPLETKONGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGGINPEKKTNCETCCPNPQSLSSGPLETKONGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
RESULT 2
AA019498
ID AA019498 standard; protein; 145 AA.
XX
AC AA019498;
XX
DT 20-DEC-2002 (first entry)
XX
DE HSI protein variant.
XX
XX HSI; variant; cancer; tumour; unigene cluster; cytostatic; metastasis;
KM EST; expressed sequence tag; colon cancer; stomach cancer; breast cancer;
KM HS169395; HS127144; HS2; HS132793; HS3.
XX
OS Unidentified.
XX
PN DE10103694-A1.
XX
PD 01-AUG-2002.
XX
PF 26-JAN-2001; 2001DE-01003694.
XX
PR 26-JAN-2001; 2001DE-01003694.
XX
PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAY.
XX
PI Brett D, Kemmerer W;
XX
DR WPI; 2002-644836/70.
XX
DR N-PSDB; AAL50100.
XX
PT Diagnosis and therapy of tumors, by determining expression rates of
PT specific expressed sequence tags of the unigene cluster, and subsequently
PT blocking their expression.
XX
PS Claim 10; Page 5; 10pp; German.
XX
CC The present invention relates to the use of expressed sequence tags
CC (ESTs), or variants, of the unigene cluster HS169395 (HS1), HS127144

CC (HS2) and/or HS132793 (HS3) for diagnosis and therapy of tumours, in
CC which their expression rates in tumour cells and/or lymph nodes are
CC determined. The EST sequences are useful as prognostic markers of
CC survival of cancer patients (high levels of EST-related mRNA are
CC associated with a poor prognosis, specifically correlated with
CC development of metastases); and for diagnosis and/or therapy of solid
CC tumours, particularly of colon, stomach and breast. The present sequence
CC is a variant of the HSI protein shown in the exemplification of the
CC invention
XX
SQ Sequence 145 AA;
Query Match 100.0%; Score 778; DB 5; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCGGSRADAIIPRYESWTRETSTWLTYYTDSADPPSAAPDSGPEAGLHSGMLEDGL 60
DB 1 MCGGSRADAIIPRYESWTRETSTWLTYYTDSADPPSAAPDSGPEAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGGINPEKKTNCETCCPNPQSLSSGPLETKONGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGGINPEKKTNCETCCPNPQSLSSGPLETKONGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
RESULT 3
ABR58646
ID ABR58646 standard; protein; 145 AA.
XX
AC ABR58646;
XX
DT 09-JUL-2003 (first entry)
XX
DE Human cancer related protein SEQ ID NO:303.
XX
XX Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
KM heart disease; atherosclerosis; endometriosis.
XX
XX Homo sapiens.
XX
OS
XX
PN MO2003025138-A2.
XX
PD 27-MAR-2003.
XX
PF 17-SEP-2002; 2002WO-US029560.
XX
PR 17-SEP-2001; 2001US-0323469P.
XX
PR 20-SEP-2001; 2001US-0323887P.
XX
PR 13-NOV-2001; 2001US-0350666P.
XX
PR 08-FEB-2002; 2002US-0355145P.
XX
PR 08-FEB-2002; 2002US-0355257P.
XX
PR 12-APR-2002; 2002US-0372246P.
XX
PA (EOSB-) EOS BIOTECHNOLOGY INC.
XX
XX Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
XX
PI Zlocnik A;
XX
XX
DR WPI; 2003-354600/33.
XX
DR N-PSDB; ACC72796.
XX
PT New genes that are up-regulated or down-regulated in cancers, useful as
PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
PT therapeutic targets for screening drugs for treating these diseases.
XX
PS Claim 12; Page 753; 767pp; English.
XX
XX
CC The present invention describes an isolated nucleic acid molecule, which
CC comprises the sequence of any of the genes that are up-regulated or down-

1.131 de c banks

CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
CC related gene nucleotide sequences which encode the proteins given in
CC ABR58521 to ABR58709. Also described: (1) determining the presence or
CC absence of a pathological cell in a patient; (2) an expression vector
CC comprising a nucleic acid molecule described above; (3) a host cell
CC comprising the vector; (4) an isolated polypeptide, which is encoded by
CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
CC of (4); (6) specifically targeting a compound to a pathological cell in a
CC patient by administering to the patient the antibody above; and (7) a
CC drug screening assay. The nucleic acid is useful as diagnostic markers or
CC therapeutic targets. In particular, the nucleic acid is useful for
CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
CC pancreas, prostate, skin and uterus), wounds, ischaemia, heart diseases,
CC atherosclerosis and endometriosis. The nucleic acid is also useful in
CC drug screening, particularly for identifying agents for treating these
CC pathologies
XX
SQ Sequence 145 AA:

Query Match 100.0%; Score 778; DB 6; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYRYESTWRETSTWLTITDSDAPPSAAPSCEAGLSGMEDEL 60
DB 1 MCGGSRADAIERYRYESTWRETSTWLTITDSDAPPSAAPSCEAGLSGMEDEL 60
QY 61 PSNGVRSRSTAPGSIPEPEKTCNCECPNPQSSSGLPTQKQGLQTTAKRDAKMPAK 120
DB 61 PSNGVRSRSTAPGSIPEPEKTCNCECPNPQSSSGLPTQKQGLQTTAKRDAKMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 4
ADCC1800 standard; protein; 145 AA.
XX
AC ADCC1800;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human novel polypeptide sequence, SEQ ID NO:1882.
XX
KW Human: diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; anti-nausea; anticoagulant; thrombolytic; vulnery;
KW anticancer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 8.
XX
XX Homo sapiens.
XX OS
XX PN WO2003029271-A2.
XX
PD 10-APR-2003.
XX
PF 24-SEP-2002; 2002WO-US030474.
XX
PR 24-SEP-2001; 2001US-0324631P.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Wang G;
PI Haley-Vicence D, Drmanac RT;
XX

DR WPI; 2003-371981/35.
DR N-PSDB; ADCC08829.
XX
PT New polynucleotide and polypeptide useful for diagnosing, preventing or
PT treating conditions such as neurodegenerative diseases, anaemias, platelet
PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX cancer.
PS Claim 20; SEQ ID NO 1882; 1185bp; English.
XX
XX The invention relates to 971 novel human cDNA sequences (ADC29919-
CC ADC30889) and the polypeptides they encode (ADC30890-ADC3160). The
CC invention also relates to nucleic acid sequences over 99% identical with
CC the novel human cDNAs. The invention additionally encompasses expression
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC1861-ADC3627) and the polypeptides encoded by the contigs (ADC2628
CC -ADC3394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 145 AA:

Query Match 100.0%; Score 778; DB 7; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYRYESTWRETSTWLTITDSDAPPSAAPSCEAGLSGMEDEL 60
DB 1 MCGGSRADAIERYRYESTWRETSTWLTITDSDAPPSAAPSCEAGLSGMEDEL 60
QY 61 PSNGVRSRSTAPGSIPEPEKTCNCECPNPQSSSGLPTQKQGLQTTAKRDAKMPAK 120
DB 61 PSNGVRSRSTAPGSIPEPEKTCNCECPNPQSSSGLPTQKQGLQTTAKRDAKMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 5
ADMA46959 standard; protein; 145 AA.
XX
AC ADMA46959;
XX
DT 03-JUN-2004 (first entry)
XX
DE Brain and Acute Leukemia, Cytoplaemic alternate protein #1.
XX
KW acute myelogenous leukemia; gene expression; BALC;
XX

KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
KM Cytoplasmic; exon.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT MISC-difference 41
PT /note= "encoded by GCS"
XX
PN WO2003040347-A2.
XX
PD 15-MAY-2003.
XX
PF 12-NOV-2002; 2002WO-US036375.
XX
PR 09-NOV-2001; 2001US-0348210P.
XX
PA (OHIS) UNIV OHIO STATE RES FOUND.
XX
PI Tanner SM, De la Chapell A;
XX
DR WPI: 2003-441564/41.
DR N-PSDB; ADM46951.
XX
PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
PT in a patient comprises assaying for the overexpression of one or more
PT BAALC transcripts in cells obtained from the patient.
XX
PS Disclosure; SEQ ID NO 17; 78pp; English.
XX
CC The invention relates to a method of characterizing acute myelogenous
CC leukemia (AML) in a patient by assaying for the overexpression of one or
CC more BAALC transcripts in cells obtained from the patient, where an
CC overexpression indicates that the patient has an aggressive form of AML.
CC The methods, kits and probes are useful for characterizing acute or
CC chronic myelogenous leukemia, or prostate cancer. They are also useful
CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
CC spliced RNA consisting of exons 1, 6 and 8.
XX
SQ Sequence 145 AA;
XX
Query Match 100.0%; Score 778; DB 7; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCCGSGRADAIIEPRRYESWTRETESWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGL 60
DB 1 MCCGSGRADAIIEPRRYESWTRETESWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGAIIPNEBKKNCECPNPSSISGPELTOKONGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGAIIPNEBKKNCECPNPSSISGPELTOKONGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145
XX
RESULT 6
ADO48475
ID ADO48475 standard; protein; 145 AA.
XX
AC ADO48475;
XX
DT 12-AUG-2004 (first entry)
XX
DE Human PTH responsive gene protein.
XX
DE PTH responsive gene; PARGB; bone-forming; bone; bone density modulation;
KM transgenic animal; osteopathic; gene therapy; osteoporosis; human.
XX
OS Homo sapiens.

XX
PN WO2004044152-A2.
XX
PD 27-MAY-2004.
XX
PF 10-NOV-2003; 2003WO-US035655.
XX
PR 12-NOV-2002; 2002US-0425532P.
XX
PA (AMHP) WYETH.
XX
PI Robinson JA, Stojanovic-Susulic V, Babić P, Murrills RJ;
XX
DR WPI: 2004-420299/39.
DR N-PSDB; ADO48474.
XX
PT New nucleic acid fragment encoding a PARGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.
XX
PS Claim 9; SEQ ID NO 4; 16pp; English.
XX
CC The invention relates to a novel PTH responsive gene (PARGB) fragment
CC encoding a polypeptide. The invention further comprises: a chimeric
CC construct comprising the isolated nucleic acid fragment operatively
CC linked to suitable regulatory sequences; a host cell transformed with the
CC chimeric construct; a vector comprising the nucleic acid fragment;
CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC obtaining a polypeptide; detecting the presence of the nucleic acid
CC fragment; an antibody that specifically binds to one or more epitopes of
CC a PARGB polypeptide; a composition for regulating bone-forming activity
CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
CC ; an agent that alters the expression of PARGB gene or polypeptide;
CC determining whether an agent alters the expression of PARGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PARGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PARGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PARGB cDNA, where upon the addition
CC of chemical inducer, transcription of PARGB gene is induced. The PARGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.
XX
SQ Sequence 145 AA;
XX
Query Match 100.0%; Score 778; DB 8; Length 145;
Best Local Similarity 100.0%; Pred. No. 4e-70;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MCCGSGRADAIIEPRRYESWTRETESWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGL 60
DB 1 MCCGSGRADAIIEPRRYESWTRETESWLTYYTDSADAPPSAAAPDSGPEAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGAIIPNEBKKNCECPNPSSISGPELTOKONGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGAIIPNEBKKNCECPNPSSISGPELTOKONGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145

```

Db      121 EVTINVTDSIQMDRSRRITKNCVN 145
|||||
RESULT 7
ID      ADM46961
XX      ADM46961 standard; protein; 180 AA.
XX      ADM46961;
XX      ADM46961;
XX      03-JUN-2004 (first entry)
XX      DT
XX      DE Brain and Acute Leukemia, Cytoplasmic alternate protein #3.
XX      KW acute myelogenous leukemia; gene expression; BAALC;
XX      KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
XX      KW Cytoplasmic; exon.
XX      OS Homo sapiens.
XX      PH Key Location/Qualifiers
XX      FT Misc-difference 41
XX      FT /note= "encoded by GCS"
XX      MO2003040347-A2.
XX      PD 15-MAY-2003.
XX      PF 12-NOV-2002; 2002WO-US036375.
XX      PR 09-NOV-2001; 2001US-0348210P.
XX      PA (OHIS ) UNIV OHIO STATE RES FOUND.
XX      PI Tanner SM, De la Chapel1 A;
XX      DR WPI: 2003-441564/41.
XX      DR N-PSDB; ADM46953.
XX      PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
XX      PT in a patient comprising assaying for the overexpression of one or more
XX      PT BAALC transcripts in cells obtained from the patient.
XX      PS Disclosure; SEQ ID NO 19; 78pp; English.
XX      CC The invention relates to a method of characterizing acute myelogenous
XX      CC leukemia (AML) in a patient by assaying for the overexpression of one or
XX      CC more BAALC transcripts in cells obtained from the patient, where an
XX      CC overexpression indicates that the patient has an aggressive form of AML.
XX      CC The methods, kits and probes are useful for characterizing acute or
XX      CC chronic myelogenous leukemia, or prostate cancer. They are also useful
XX      CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
XX      CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
XX      CC spliced RNA consisting of exons 1, 6 and 8.
XX      SO Sequence 180 AA:
XX
Query Match 96.5%; Score 750.5; DB 7; Length 180;
Best Local Similarity 80.6%; Pred. No. 3.1e-67;
Matches 145; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

```

```

RESULT 8
ID      ADO48479
XX      ADO48479 standard; protein; 145 AA.
XX      ADO48479;
XX      ADO48479;
XX      12-AUG-2004 (first entry)
XX      DT
XX      DE Mouse PTH responsive gene protein.
XX      KW PTH responsive gene; PAlGB; bone-forming; bone; bone density modulation;
XX      KW transgenic animal; osteopathic; gene therapy; osteoporosis; mouse;
XX      KW murine.
XX      OS Mus sp.
XX      PH MO200404152-A2.
XX      PN 27-MAY-2004.
XX      PD 10-NOV-2003; 2003WO-US035655.
XX      PF 12-NOV-2002; 2002US-0425532P.
XX      PR (AMHP ) WYETH.
XX      PA Robinson JA, Stojanovic-Susulic V, Babij P, Murrills RJ;
XX      PI WPI: 2004-420299/39.
XX      DR N-PSDB; ADO48478.
XX      PT New nucleic acid fragment encoding a PAlGB polypeptide, useful in
XX      PT preparing a composition for diagnosing, treating or preventing bone
XX      PT related disorders, e.g., osteoporosis.
XX      PS Claim 9; SEQ ID NO 8; 169pp; English.
XX
The invention relates to a novel PTH responsive gene (PAlGB) fragment
encoding a polypeptide. The invention further comprises: a chimeric
construct comprising the isolated nucleic acid fragment operatively
linked to suitable regulatory sequences; a host cell transformed with the
chimeric construct; a vector comprising the nucleic acid fragment;
obtaining a nucleic acid fragment encoding the polypeptide; a method for
obtaining a polypeptide; detecting the presence of the nucleic acid
fragment; an antibody that specifically binds to one or more epitopes of
a PAlGB polypeptide; a composition for regulating bone-forming activity
in a mammal comprising the nucleic acid fragment, polypeptide or antibody
; an agent that alters the expression of PAlGB gene or polypeptide;
determining whether an agent alters the expression of PAlGB mRNA;
screening agents for effectiveness in altering expression of the nucleic
acid fragment; screening for agents useful for treating bone related
disorders; evaluating the efficacy of a treatment of a bone related
disorder in a subject; identifying polypeptides capable of binding to
PAlGB; monitoring the effectiveness of treatment of a subject with a bone
related agent; a transgenic animal comprising the DNA, an animal model
for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PAlGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PAlGB cDNA, where upon the addition
CC of chemical inducer, transactivation of PAlGB gene is induced. The PAlGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.

```

XX Sequence 145 AA;
 SQ Query Match 83.4%; Score 649; DB 8; Length 145;
 Best Local Similarity 83.4%; Pred. No. 4e-57;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 QY 1 MCGGSRADAIERPRYESWTRETSTWLTYSDDAPSPAAPDSGPEAGLHSGMLBDGL 60
 DB 1 MCGGSRADAIERPRYESWTRETSTWLTYSDDAPSPAAPDSGPEAGLHSGMLBDGL 60
 QY 61 PSNGVPRSTAPGGINPEKKTNCETQCPNPSGLSGPLTQKONGLOTTEAKRDAKMPAK 120
 DB 61 SSNGVLRPAAPGGINPEKKNCGTQCPNPSGLSGPLTQKONGLOTTEAKRDAKMSAR 120
 QY 121 EYTIWVTSIQMDRSRRITKNCVN 145
 DB 121 EVAIWNTENIRQMDRSKRVTKNCIN 145

RESULT 9

AD048473
 ID AD048473 standard; protein; 145 AA.

AC AD048473;
 XX

DT 12-AUG-2004 (first entry)
 XX

DE Rat PTH responsive gene protein.
 XX

KW PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
 transgenic animal; osteopathic; gene therapy; osteoporosis; rat.
 XX

OS Rattus sp.
 XX

PN WO2004044152-A2.
 XX

PD 27-MAY-2004.
 XX

PF 10-NOV-2003; 2003WO-US035655.
 XX

PR 12-NOV-2002; 2002US-0425532P.
 XX

PA (AMHP) WYETH.
 XX

PI Robinson JA, Stojanovic-Susulic V, Babij P, Murrills RJ;
 XX

DR WPI; 2004-420239/39.
 XX

DR N-PSDB; AD048472.
 XX

PT New nucleic acid fragment encoding a PAIGB polypeptide, useful in
 preparing a composition for diagnosing, treating or preventing bone
 related disorders, e.g., osteoporosis.
 XX

PS Claim 9; SEQ ID NO 2; 169pp; English.
 XX

CC The invention relates to a novel PTH responsive gene (PAIGB) fragment
 encoding a polypeptide. The invention further comprises: a chimeric
 construct comprising the isolated nucleic acid fragment operatively
 linked to suitable regulatory sequences; a host cell transfected with the
 chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 obtaining a polypeptide; detecting the presence of the nucleic acid
 fragment; an antibody that specifically binds to one or more epitopes of
 a PAIGB polypeptide; a composition for regulating bone-forming activity
 in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 CC ; an agent that alters the expression of PAIGB gene or polypeptide;
 CC determining whether an agent alters the expression of PAIGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 acid fragment; screening for agents useful for treating bone related
 disorders; evaluating the efficacy of a treatment of a bone related
 disorder in a subject; identifying polypeptides capable of binding to
 CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone

CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain
 CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.
 XX

XX Sequence 145 AA;

Query Match 82.9%; Score 645; DB 8; Length 145;
 Best Local Similarity 83.4%; Pred. No. 1e-56;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRETSTWLTYSDDAPSPAAPDSGPEAGLHSGMLBDGL 60
 DB 1 MCGGSRADAIERPRYESWTRETSTWLTYSDDAPSPAAPDSGPEAGLHSGMLBDGL 60

QY 61 PSNGVPRSTAPGGINPEKKTNCETQCPNPSGLSGPLTQKONGLOTTEAKRDAKMPAK 120
 DB 61 SSNGVLRPAAPGGINPEKKNCGTQCPNPSGLSGPLTQKONGLOTTEAKRDAKMSAR 120

QY 121 EYTIWVTSIQMDRSRRITKNCVN 145
 DB 121 EVAIWNTENIRQMDRSKRVTKNCIN 145

RESULT 10

AD046963
 ID AD046963 standard; protein; 149 AA.

AC AD046963;
 XX

DT 03-JUN-2004 (first entry)
 XX

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #5.
 XX

KW acute myelogenous leukemia; gene expression; BALC;
 chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 XX Cytoplasmic; exon.

OS Homo sapiens.
 XX

PN WO2003040347-A2.
 XX

PD 15-MAY-2003.
 XX

PF 12-NOV-2002; 2002WO-US036375.
 XX

PR 09-NOV-2001; 2001US-0348210P.
 XX

PA (OHIS) UNIV OHIO STATE RES FOUND.
 XX

PI Tanner SM, De La Chapell A;
 XX

DR WPI; 2003-441564/41.
 XX

DR N-PSDB; AD046955.
 XX

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprises assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.
 XX
 XX disclosure; SEQ ID NO 21; 78pp; English.
 XX
 CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.
 XX
 SQ Sequence 149 AA;
 XX
 Query Match 73.2%; Score 569.5; DB 7; Length 149;
 Best Local Similarity 75.7%; Pred. No. 4,2e-49;
 Matches 109; Conservative 0; Mismatches 0; Indels 35; Gaps 1;
 XX
 QY 1 MGGGSGRADAIERRYESWTRETESTWLTGTTSDAPPSPAAPDSGEGAGLHS----- 53
 DB 1 MGGGSGRADAIERRYESWTRETESTWLTGTTSDAPPSPAAPDSGEGAGLHSEAEKXS 60
 QY 54 -----GMLDGLPSNGVPRSTAGGIPNPEKTKNCET 85
 DB 61 KIKAPDVSDEGLFSAKMAPLAVFSGMLDGLPSNGVPRSTAGGIPNPEKTKNCET 120
 QY 86 QCPNPOSLSGGPLTOKONGLOTTE 109
 DB 121 QCPNPOSLSGGPLTOKONGLOTTE 144
 XX
 RESULT 11
 ADQ81902
 ID ADQ81902 standard; protein; 92 AA.
 AC ADQ81902;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Human dihydrogenase 10.12.
 XX
 KW Human; enzyme; dihydrogenase 10.12; malignant tumour; inflammation;
 KW immunological disease; haemopathy; HIV infection.
 XX
 OS Homo sapiens.
 XX
 PN CN1344798-A.
 XX
 PD 17-APR-2002.
 XX
 PF 29-SEP-2000; 2000CN-00125495.
 XX
 PR 29-SEP-2000; 2000CN-00125495.
 XX
 PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI, 2002-509506/55.
 DR N-PSDB; ADQ81901.
 XX
 PT New polypeptide human dihydrogenase 10.12 and polynucleotides encoding this
 PT polypeptide, useful for treating various diseases, such as malignant
 PT tumors, inflammations, immunological diseases, hemopathy and HIV
 PT infection.
 XX
 PS Claim 1; SEQ ID NO 2; 33pp; Chinese.
 XX
 CC The present invention discloses a new kind of polypeptide, human

CC dihydrogenase 10.12, polynucleotides encoding this polypeptide, a DNA
 CC recombination process to produce the polypeptide and antagonist against
 CC the polypeptide. The present invention also discloses the method of
 CC applying the polypeptide in treating various diseases, such as malignant
 CC tumors, inflammations, immunological diseases, haemopathy and HIV
 CC infection. The present sequence is the human dihydrogenase 10.12.
 XX
 SQ Sequence 92 AA;
 XX
 Query Match 52.1%; Score 405.5; DB 5; Length 92;
 Best Local Similarity 79.2%; Pred. No. 7.5e-33;
 Matches 80; Conservative 0; Mismatches 2; Indels 19; Gaps 1;
 XX
 QY 45 GEGAGLHSGMLEDGLPSNGVPRSTAGGIPNPEKTKNCETCCPNPOSLSGGPLTOKONG 104
 DB 11 GPEFG-----NAPGIPNPEKTKNCETCCPNPOSLSGGPLTOKONG 51
 QY 105 LQTEAKRDAKEMPAKEVTINTVDSIQMDRSRRTKNCVN 145
 DB 52 LQTEAKRDAKEMPAKEVTINTVDSIQMDRSRRTKNCVN 92
 XX
 RESULT 12
 ADM46962
 ID ADM46962 standard; protein; 73 AA.
 AC ADM46962;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Brain and Acute Leukemia, Cytoplasmic alternate protein #4.
 XX
 KW acute myelogenous leukemia; gene expression; BAALC;
 KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 KW Cytoplasmic; exon.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 41
 XX
 FT /note= "encoded by GCS"
 XX
 PN MO2003040347-A2.
 XX
 PD 15-MAY-2003.
 XX
 PF 12-NOV-2002; 2002WO-US036375.
 XX
 PR 09-NOV-2001; 2001US-0348210P.
 XX
 PA (OHIS) UNIV OHIO STATE RES FOUND.
 XX
 PI Tanner SM, De la Chapell A;
 XX
 DR WPI: 2003-441564/41.
 DR N-PSDB; ADM46954.
 XX
 PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprises assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.
 XX
 PS disclosure; SEQ ID NO 20; 78pp; English.
 XX
 CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.
 XX

SQL Sequence 73 AA;

Query Match 39.7%; Score 309; DB 7; Length 73;
Best Local Similarity 96.6%; Pred. No. 2, 9e-23;
Matches 56; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIEPRYESTWRTETSTWLTFTYDSDAPPSPAAPDGPAGGLHSGMLD 58
1 MCGGSRADAIIEPRYESTWRTETSTWLTFTYDSDAPPSPAAPDGPAGGLHSGCLEE 58

RESULT 13

ADM46960
ID ADM46960 standard; protein; 54 AA.

AC ADM46960;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #2.

KM acute myelogenous leukemia; gene expression; BALC;
KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
KM Cytoplasmic; exon.

OS Homo sapiens.

Key Location/Qualifiers

FT Misc-difference 41 /note="encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

DR WPI; 2003-441564/41.

DR N-PSDB; ADM46952.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
PT in a patient comprising assaying for the overexpression of one or more
PT BALC transcripts in cells obtained from the patient.

PS Disclosure: SEQ ID NO 18; 78pp; English.

CC The invention relates to a method of characterizing acute myelogenous
CC leukemia (AML) in a patient by assaying for the overexpression of one or
CC more BALC transcripts in cells obtained from the patient, where an
CC overexpression indicates that the patient has an aggressive form of AML.
CC The methods, kits and probes are useful for characterizing acute or
CC chronic myelogenous leukemia, or prostate cancer. They are also useful
CC for detecting BALC overexpression. This sequence corresponds to a BALC
CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
CC spliced RNA consisting of exons 1, 6 and 8.

XX Sequence 54 AA;

Query Match 38.4%; Score 299; DB 7; Length 54;
Best Local Similarity 100.0%; Pred. No. 2e-22;

Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIEPRYESTWRTETSTWLTFTYDSDAPPSPAAPDGPAGGLHSG 54
1 MCGGSRADAIIEPRYESTWRTETSTWLTFTYDSDAPPSPAAPDGPAGGLHSG 54

RESULT 14

ID ADO48477
ID ADO48477 standard; protein; 54 AA.

AC ADO48477;

DT 12-AUG-2004 (first entry)

DE Human PTH responsive gene protein exon 2 splice variant.

KM PTH responsive gene; PARGB; bone-forming; bone; bone density modulation;
KM transgenic animal; osteopathic; gene therapy; osteoporosis; human.

OS Homo sapiens.

PN WO2004044152-A2.

PD 27-MAY-2004.

PF 10-NOV-2003; 2003WO-US035655.

PR 12-NOV-2002; 2002US-0425532P.

PA (AMHP) WYETH.

PI Robinson UA, Stojanovic-Susulic V, Babij P, Murrills RJ;

DR WPI; 2004-420299/39.

DR N-PSDB; ADO48476.

PT New nucleic acid fragment encoding a PARGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.

PS Claim 9; SEQ ID NO 6; 16pp; English.

CC The invention relates to a novel PTH responsive gene (PARGB) fragment
CC encoding a polypeptide. The invention further comprises: a chimeric
CC construct comprising the isolated nucleic acid fragment operatively
CC linked to suitable regulatory sequences; a host cell transformed with the
CC chimeric construct; a vector comprising the nucleic acid fragment;
CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC obtaining a polypeptide; detecting the presence of the nucleic acid
CC fragment; an antibody that specifically binds to one or more epitopes of
CC a PARGB polypeptide; a composition for regulating bone-forming activity
CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
CC; an agent that alters the expression of PARGB gene or polypeptide;
CC determining whether an agent alters the expression of PARGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PARGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PARGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PARGB cDNA, where upon the addition
CC of chemical inducer, transcription of PARGB gene is induced. The PARGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.

XX Sequence 54 AA;

Job time : 105.667 secs

Query Match 38.4%; Score 299; DB 8; Length 54;
 Best Local Similarity 100.0%; Pred. No. 2e-22;
 Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGCGSRADAIEPRYESWTRETESTWLTYTDSAPPSAAPDGGPEAGGHS 54
 |||||||
 DB 1 MCGCGSRADAIEPRYESWTRETESTWLTYTDSAPPSAAPDGGPEAGGHS 54

RESULT 15

ADM46964
 ID ADM46964 standard; protein; 80 AA.

XX
 AC ADM46964;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #6.

KM acute myelogenous leukemia; gene expression; BAALC;
 KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 XX Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN MO2003040347-A2.

PD 15-MAY-2003.

XX 12-NOV-2002; 2002WC-US036375.

XX 09-NOV-2001; 2001US-0348210P.

XX (OHIS) UNIV OHIO STATE RES FOUND.

XX Tanner SM, De la Chapell A;

XX WPI; 2003-441564/41.

DR N-PSDB; ADM46956.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprises assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 22; 78bp; English.

XX
 CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.

XX
 SQ Sequence 80 AA;

Query Match 37.7%; Score 293; DB 7; Length 80;

Best Local Similarity 100.0%; Pred. No. 1.4e-21;
 Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGCGSRADAIEPRYESWTRETESTWLTYTDSAPPSAAPDGGPEAGGHS 53
 |||||||
 DB 1 MCGCGSRADAIEPRYESWTRETESTWLTYTDSAPPSAAPDGGPEAGGHS 53

Search completed: April 12, 2005, 15:37:18

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

CM protein - protein search, using sw model

Run on: April 12, 2005, 15:15:13 ; Search time 103.667 Seconds

(without alignment)
540,967 Million cell updates/sec

Title: US-10-705-716a-2

Perfect score: 767
Sequence: 1 MCGGSGRADAIERYESWT.....VTENIRQMDKSKVTNKCN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :

A_GeneSeq_16Dec04:.*
1: geneseqp19808:.*
2: geneseqp19808:.*
3: geneseqp20008:.*
4: geneseqp20018:.*
5: geneseqp20028:.*
6: geneseqp20038:.*
7: geneseqp20038:.*
8: geneseqp20048:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	767	100.0	145	AD048473	AD048473 Rat PTH r
2	746	97.3	145	AD048479	AD048479 Mouse PTH
3	645	84.1	145	AB095018	AB095018 Human pro
4	645	84.1	145	AA019498	AA019498 HSI prote
5	645	84.1	145	ABR58646	ABR58646 Human can
6	645	84.1	145	AD031800	AD031800 Human nov
7	645	84.1	145	AD046959	AD046959 Brain and
8	645	84.1	145	AD048475	AD048475 Human PTH
9	617.5	80.5	180	AD048475	AD048475 Human PTH
10	473.5	61.7	149	AD048475	AD048475 Human PTH
11	335.5	43.7	92	AD081902	AD081902 Human and
12	294	38.3	54	AD048481	AD048481 Mouse PTH
13	288	37.5	73	AD046962	AD046962 Brain and
14	278	36.2	54	AD046960	AD046960 Brain and
15	278	36.2	54	AD048477	AD048477 Human PTH
16	272	35.5	80	AD048477	AD048477 Human PTH
17	140	18.3	25	AD046964	AD046964 Brain and
18	99	12.9	18	AD046977	AD046977 Brain and
19	92.5	12.1	15	AD046977	AD046977 Brain and
20	89	11.6	16	AD048482	AD048482 PTH respo
21	87.5	11.4	260	AB098989	AB098989 Novel hum
22	87.5	11.4	592	ABG15607	ABG15607 Novel hum
23	84	11.0	718	ABU17344	ABU17344 Protein e
24	84	11.0	726	ADA36828	ADA36828 Actinotro
25	83.5	10.9	700	AAE25052	AAE25052 Stenotro

26	83	10.8	1433	5	ABP35624	ABP35624 Fungal ZB
27	82	10.7	339	6	AAE14866	AAE14866 S. clavul
28	81.5	10.6	600	7	ADJ70139	ADJ70139 Human hea
29	81.5	10.6	852	4	AA040296	AA040296 Human pol
30	81.5	10.6	872	4	AA040295	AA040295 Human pol
31	81.5	10.6	886	4	AA042081	AA042081 Human pol
32	81.5	10.6	886	4	AA042082	AA042082 Human pol
33	81.5	10.6	950	8	AD089824	AD089824 Antagonis
34	81.5	10.6	974	4	AAU15506	AAU15506 Novel hum
35	80.5	10.5	275	4	ABG21379	ABG21379 Novel hum
36	80.5	10.5	275	4	ABG15431	ABG15431 Novel hum
37	80.5	10.5	1100	4	AA084930	AA084930 Shrimp wh
38	79.5	10.4	286	2	AA000097	AA000097 Enterococ
39	79.5	10.4	286	5	ABP43316	ABP43316 E. faecali
40	79.5	10.4	286	6	ABU88344	ABU88344 E. faecali
41	79.5	10.4	286	6	ABU13595	ABU13595 Enterococ
42	79.5	10.4	305	2	AA000096	AA000096 Enterococ
43	79.5	10.4	305	5	ABP43315	ABP43315 E. faecali
44	79.5	10.4	305	6	ABU88343	ABU88343 E. faecali
45	79.5	10.4	305	6	ABU13594	ABU13594 Enterococ

ALIGNMENTS

RESULT 1
AD048473 standard; protein, 145 AA.
XX
XX AD048473;
XX
AC
XX
DT 12-AUG-2004 (first entry)
XX
DE Rat PTH responsive gene protein.
XX
KM PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
KM transgenic animal; osteopathic; gene therapy; osteoporosis; rat.
XX
XX Rattus sp.
OS
XX
XX WO2004044152-A2.
PN
XX
PD 27-MAY-2004.
XX
XX 10-NOV-2003; 2003WO-US035655.
XX
XX 12-NOV-2002; 2002US-0425532P.
XX
XX (AMHP) WYETH.
PA
XX Robinson JA, Stojanovic-Suenilic V, Babić P, Murrills RJ;
XX WPI; 2004-420239/39.
DR N-PSDB; AD048472.
XX
XX New nucleic acid fragment encoding a PAIGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.
XX
XX Claim 9; SEQ ID NO 2; 169pp; English.
PS
XX The invention relates to a novel PTH responsive gene (PAIGB) fragment
XX encoding a polypeptide. The invention further comprises: a chimeric
XX construct comprising the isolated nucleic acid fragment operatively
XX linked to suitable regulatory sequences; a host cell transformed with the
XX chimeric construct; a vector comprising the nucleic acid fragment;
XX obtaining a nucleic acid fragment encoding the polypeptide; a method for
XX obtaining a polypeptide; detecting the presence of the nucleic acid
XX fragment; an antibody that specifically binds to one or more epitopes of
XX a PAIGB polypeptide; a composition for regulating bone-forming activity
XX in a mammal comprising the nucleic acid fragment, polypeptide or antibody
XX; an agent that alters the expression of PAIGB gene or polypeptide;
XX determining whether an agent alters the expression of PAIGB mRNA;

CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related
 CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain
 CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.

CC Sequence 145 AA;

Query Match 100.0%; Score 767; DB 8; Length 145;

Best Local Similarity 100.0%; Pred. No. 4, 1e-76; Indels 0; Gaps 0;

Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRETESTTWLTYSDDLPSAATDGPAGGLHAGVLEDP 60
 |||||
 Db 1 MCGGSRADAIERPRYESWTRETESTTWLTYSDDLPSAATDGPAGGLHAGVLEDP 60

QY 61 SSGNGVLRPAAPGCIANPEKMNCGTQCPNSQSLSSGPILOKONGMLTTAKRDAKMSAR 120
 |||||
 Db 61 SSGNGVLRPAAPGCIANPEKMNCGTQCPNSQSLSSGPILOKONGMLTTAKRDAKMSAR 120

QY 121 EVAISTENIRQMDRSKRVTKNCIN 145
 |||||
 Db 121 EVAISTENIRQMDRSKRVTKNCIN 145

RESULT 2

ADO48479 standard; protein; 145 AA.

XX ADO48479;

XX 12-AUG-2004 (first entry)

XX Mouse PTH responsive gene protein.

XX PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;

XX transgenic animal; osteopathic; gene therapy; osteoporosis; mouse;

XX murine.

XX Mus sp.

XX WO2004044152-A2.

XX 27-MAY-2004.

XX 10-NOV-2003; 2003WO-US035655.

XX 12-NOV-2002; 2002US-0425532P.

XX (AMHP) WYETH.

XX Robinson JA, Stojanovic-Susulic V, Babić P, Murrills RJ;

XX WPI; 2004-420299/39.

XX N-PSDB; ADO48478.

XX New nucleic acid fragment encoding a PAIGB polypeptide, useful in
 PT preparing a composition for diagnosing, treating or preventing bone
 PT related disorders, e.g., osteoporosis.

XX Claim 9; SEQ ID NO 8; 169pp; English.

XX The invention relates to a novel PTH responsive gene (PAIGB) fragment
 CC encoding a polypeptide. The invention further comprises: a chimeric
 CC construct comprising the isolated nucleic acid fragment operatively
 CC linked to suitable regulatory sequences; a host cell transformed with the
 CC chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 CC obtaining a polypeptide; detecting the presence of the nucleic acid
 CC fragment; an antibody that specifically binds to one or more epitopes of
 CC a PAIGB polypeptide; a composition for regulating bone-forming activity
 CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 CC ; an agent that alters the expression of PAIGB gene or polypeptide;
 CC determining whether an agent alters the expression of PAIGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related
 CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain
 CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.

XX Sequence 145 AA;

Query Match 97.3%; Score 746; DB 8; Length 145;

Best Local Similarity 97.2%; Pred. No. 8, 4e-74; Indels 0; Gaps 0;

Matches 141; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESWTRETESTTWLTYSDDLPSAATDGPAGGLHAGVLEDP 60
 |||||
 Db 1 MCGGSRADAIERPRYESWTRETESTTWLTYSDDLPSAATDGPAGGLHAGVLEDP 60

QY 61 SSGNGVLRPAAPGCIANPEKMNCGTQCPNSQSLSSGPILOKONGMLTTAKRDAKMSAR 120
 |||||
 Db 61 SSGNGVLRPAAPGCIANPEKMNCGTQCPNSQSLSSGPILOKONGMLTTAKRDAKMSAR 120

QY 121 EVAISTENIRQMDRSKRVTKNCIN 145
 |||||
 Db 121 EVAISTENIRQMDRSKRVTKNCIN 145

RESULT 3

AAB95018 standard; protein; 145 AA.

XX AAB95018;

XX 26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:16726.

XX Human; primer; detection; diagnosis; antisense therapy; gene therapy.

```

XX OS Homo sapiens.
XX XX EPI074617-A2.
XX FM EPI074617-A2.
XX PD 07-FEB-2001.
XX PE 28-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99UP-00248036.
XX PR 27-AUG-1999; 99UP-00300253.
XX PR 11-JAN-2000; 2000UP-00118776.
XX PR 02-MAY-2000; 2000UP-0018367.
XX PR 09-JUN-2000; 2000UP-00241899.
XX XX
XX PA (HELI-) HELIX RES INST.
XX PI Ota T, Inogai T, Miehikawa T, Hayashi K, Saito K, Yamamoto J;
XX PI Ishii S, Sugiyama T, Makamatsu A, Nagai K, Otsuki T;
XX XX
XX DR WPI, 2001-318749/34.
XX XX
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS
XX PS Claim 8; SEQ ID NO 16726; 2537pp + Sequence Listing; English.
XX CC
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-OT primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence, where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence/3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAH92446 to AAB95893
XX CC represent human amino acid sequences; and AAH18629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX XX
XX BQ Sequence 145 AA;
XX
XX Query Match 84.1%; Score 645; DB 4; Length 145;
XX Best Local Similarity 83.4%; Pred. No. 1, 2e-62;
XX Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0
XX
OY 1 MCCGGSRADAIIEPRRYESWTRETSFWLYTTSDALPSAAATDSGPGAGLHAGVLEDCP 60
DB 1 MCCGGSRADAIIEPRRYESWTRETSFWLYTTSDADPPSAAAPDSGPGAGLHGMLEDGL 60
OY 61 SNGVLRPAAPGAIAPPEKKMNGTCQPNSSQSLSSGPTLQKONGLWTTTEAKRPAXKMSAR 120
DB 61 PENSNGPRSTRAPGIIPEPEKKTCTCPNPQSSLPPLQKONGLOTTEAKRKDAKMPAK 120
OY 121 EVAISVTENIRQMDRSKRVTKCIN 145
DB 121 EYINVTDSIQMDRSRRITTKCVN 145
XX
XX RESULT 4
XX MA019498

```

XX DE Human cancer related protein SEQ ID NO:303.
 XX XX
 XX KW Human; cancer; diagnosis; screening; modulator; leukaemia; ischaemia;
 XX KW heart disease; atherosclerosis; endometriosis.
 XX OS Homo sapiens.
 XX PN WO2003025138-A2.
 XX PD 27-MAR-2003.
 XX PF 17-SEP-2002; 2002WO-US029560.
 XX PR 17-SEP-2001; 2001US-0323469P.
 XX PR 20-SEP-2001; 2001US-0323487P.
 XX PR 13-NOV-2001; 2001US-0350666P.
 XX PR 08-FEB-2002; 2002US-0355145P.
 XX PR 08-FEB-2002; 2002US-0355257P.
 XX PR 12-APR-2002; 2002US-0372246P.
 XX PA (BOSB-) EOS BIOTECHNOLOGY INC.
 XX PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE,
 XX PI Zlotnick A;
 XX DR WPI; 2003-354600/33.
 XX DR N-PSDB; ACC22796.
 XX PT New genes that are up-regulated or down-regulated in cancers, useful as
 XX PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 XX PT therapeutic targets for screening drugs for treating these diseases.
 XX PS Claim 12; Page 753; 767pp; English.
 CC The present invention describes an isolated nucleic acid molecule, which
 CC comprises the sequence of any of the genes that are up-regulated or down-
 CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
 CC related gene nucleotide sequences which encode the proteins given in
 CC ABR8521 to ABR85709. Also described: (1) determining the presence or
 CC absence of a pathological cell in a patient; (2) an expression vector
 CC comprising a nucleic acid molecule described above; (3) a host cell
 CC comprising the vector; (4) an isolated polypeptide, which is encoded by
 CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
 CC of (4); (6) specifically targeting a compound to a pathological cell in a
 CC patient by administering to the patient the antibody above; and (7) a
 CC drug screening assay. The nucleic acid is useful as diagnostic markers or
 CC therapeutic targets. In particular, the nucleic acid is useful for
 CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
 CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
 CC pancreas, prostate, skin and uterus), wounds, ischemia, heart diseases,
 CC atherosclerosis and endometriosis. The nucleic acid is also useful in
 CC drug screening, particularly for identifying agents for treating these
 CC pathologies
 XX SO Sequence 145 AA;
 Query Match 84.1%; Score 645; DB 6; Length 145;
 Best Local Similarity 83.4%; Pred. No. 1.2e-62;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
 QY 1 MCCGSGRADAIIPRYESWTRETSTWLTYYTSDALPSAATDSCGPEAGLHAGVLEDDP 60
 DB 1 MCCGSGRADAIIPRYESWTRETSTWLTYYTSDAPPSAAAPDSGPEAGLHAGVLEDDP 60
 QY 61 SSNGVLRPAAPGGINANPEKKNACGTCCPSQSISGPTLQKONGIWTTEAKRDAXKMSAR 120
 DB 61 PSNGVLRPAAPGGINANPEKKNACGTCCPSQSISGPTLQKONGIWTTEAKRDAXKMSAR 120
 QY 121 EVAISVTENIRQMSKRYTKNCIN 145
 DB 121 EVTINVTDSIQQMDRSRRITKNCVN 145

RESULT 6
 ID ADC31800
 ID ADC31800 standard; protein; 145 AA.
 XX AC ADC31800;
 XX AC ADC31800;
 XX DT 18-DEC-2003 (first entry)
 XX DE Human novel polypeptide sequence, SEQ ID NO:1882.
 XX KW Human; diagnostic; drug screening; forensics; gene mapping;
 KW biologically assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vlnnary;
 KW anticler; osteopathic; immunosuppressive; antiinflammatory; cyrostatic;
 KW gene therapy; chromosome 8.
 XX OS Homo sapiens.
 XX PN WO2003029271-A2.
 XX PD 10-APR-2003.
 XX PR 24-SEP-2002; 2002WO-US030474.
 XX PR 24-SEP-2001; 2001US-0324631P.
 XX PA (HYSE-) HYSEQ INC.
 XX PI Tang TY, Zhang J, Ren F, Yue AJ, Zhao QH, Wang J, Wehrman T;
 XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 XX PI Haley-Vicente D, Dmatnac RT;
 XX DR WPI; 2003-371981/35.
 XX DR N-PSDB; ADC30829.
 XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 XX PT treating conditions such as neurodegenerative diseases, anemia, platelet
 XX PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 XX PS cancer.
 XX PS Claim 20; SEQ ID NO 1882; 1185pp; English.
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC coding sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32677) and the polypeptides encoded by the cDNAs (ADC32628
 CC -ADC33994). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides

CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human polypeptide sequence of the invention. Note: The sequence
CC data for this patent did not form part of the printed specification, but
CC was obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 145 AA;

Query Match 84.1%; Score 645; DB 7; Length 145;
Best Local Similarity 83.4%; Pred. No. 1.2e-62;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MGGGSGRADAIERPRYESWTRETSTWLTYYTSDALPSAATDSGPEAGGLHAGVLEDGP 60
DB 1 MGGGSGRADAIERPRYESWTRETSTWLTYYTSDALPSAAPDSGPEAGGLHSGMLEDGL 60
QY 61 SSSGVLRLPAPGGIAPPEKKNCGTCCPNSQSLSGSPFLTKQKGLMTTEAKRKAKMSAR 120
DB 61 PSNGVPRSTAPGGIPNPEKKTNCETCCPNQSSLSGFLTKQKGLQTTTEAKRKAKMPAK 120
QY 121 EVAISTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 7
ADM46959 standard; protein; 145 AA.

XX ADM46959;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #1.

XX acute myelogenous leukemia; gene expression; BAAIC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

KM Cytoplasmic; exon.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Misc-difference 41 /note="encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapel A;

DR WPI; 2003-441564/41.

DR N-PSDB; ADM46951.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer

XX in a patient comprises assaying for the overexpression of one or more

XX BAAIC transcripts in cells obtained from the patient.

XX Disclosure; SEQ ID NO 17; 78bp; English.

XX The invention relates to a method of characterizing acute myelogenous

XX leukemia (AML) in a patient by assaying for the overexpression of one or

XX more BAAIC transcripts in cells obtained from the patient, where an

XX overexpression indicates that the patient has an aggressive form of AML.

XX The methods, kits and probes are useful for characterizing acute or

XX chronic myelogenous leukemia, or prostate cancer. They are also useful

CC for detecting BAAIC overexpression. This sequence corresponds to a BAAIC

CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative

CC spliced RNA consisting of exons 1, 6 and 8.

XX Sequence 145 AA;

Query Match 84.1%; Score 645; DB 7; Length 145;
Best Local Similarity 83.4%; Pred. No. 1.2e-62;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MGGGSGRADAIERPRYESWTRETSTWLTYYTSDALPSAATDSGPEAGGLHAGVLEDGP 60
DB 1 MGGGSGRADAIERPRYESWTRETSTWLTYYTSDALPSAAPDSGPEAGGLHSGMLEDGL 60
QY 61 SSSGVLRLPAPGGIAPPEKKNCGTCCPNSQSLSGSPFLTKQKGLMTTEAKRKAKMSAR 120
DB 61 PSNGVPRSTAPGGIPNPEKKTNCETCCPNQSSLSGFLTKQKGLQTTTEAKRKAKMPAK 120
QY 121 EVAISTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 8
ADO48475 standard; protein; 145 AA.

XX ADO48475;

DT 12-AUG-2004 (first entry)

DE Human PTH responsive gene protein.

KW PTH responsive gene; PARG; bone-forming; bone; bone density modulation;

KW transgenic animal; osteopathic; gene therapy; osteoporosis; human.

XX Homo sapiens.

PN WO2004041152-A2.

PD 27-MAY-2004.

PF 10-NOV-2003; 2003WO-US035655.

PR 12-NOV-2002; 2002US-0425532P.

PA (AMHP) WYETH.

PI Robinson JA, Stojanovic-Suenic V, Babij P, Murrills RJ;

DR WPI; 2004-420299/39.

DR N-PSDB; ADO48474.

PT New nucleic acid fragment encoding a PARG polypeptide, useful in

XX preparing a composition for diagnosing, treating or preventing bone

XX related disorders, e.g., osteoporosis.

XX Claim 9; SEQ ID NO 4; 169bp; English.

XX The invention relates to a novel PTH responsive gene (PARG) fragment

XX encoding a polypeptide. The invention further comprises a chimeric

XX construct comprising the isolated nucleic acid fragment operatively

XX linked to suitable regulatory sequences; a host cell transformed with the

XX chimeric construct; a vector comprising the nucleic acid fragment;

XX obtaining a polypeptide; detecting the presence of the nucleic acid

XX fragment; an antibody that specifically binds to one or more epitopes of

XX a PARG polypeptide; a composition for regulating bone-forming activity

XX in a mammal comprising the nucleic acid fragment, polypeptide or antibody

XX ; an agent that alters the expression of PARG gene or polypeptide;

XX determining whether an agent alters the expression of PARG mRNA;

XX screening agents for effectiveness in altering expression of the nucleic

XX acid fragment; screening for agents useful for treating bone related

CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA, an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity, and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.

SQ Sequence 145 AA;

Query Match

84.1%; Score 645; DB 8; Length 145;

Best Local Similarity 83.4%; Pred. No. 1.2e-62; Mismatches 14; Indels 0; Gaps 0;

Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETSWLTYYTDSALPSAANTDSGPAGGLHAGVLEDP 60
|||||
1 MCGGSRADAIIPRYESWTRETSWLTYYTDSALPSAANTDSGPAGGLHAGVLEDP 60
Db 1 MCGGSRADAIIPRYESWTRETSWLTYYTDSALPSAANTDSGPAGGLHAGVLEDP 60
QY 61 SSGVVRPAAPGSIAPPEKKNCGTCCPNSSGSLSGPLTQKNGLTWTEAKRDKKMSAR 120
|||||
61 PSNGVVRPAAPGSIAPPEKKNCGTCCPNSSGSLSGPLTQKNGLTWTEAKRDKKMSAR 120
Db 61 PSNGVVRPAAPGSIAPPEKKNCGTCCPNSSGSLSGPLTQKNGLTWTEAKRDKKMSAR 120
QY 121 EVAISVTENIRQMDRSKRVTKNCIN 145
|||
121 EVAISVTENIRQMDRSKRVTKNCIN 145
Db 121 EVAISVTENIRQMDRSKRVTKNCIN 145

RESULT 9

ADM46961 standard; protein; 180 AA.

XX ADM46961;

XX 03-JUN-2004 (first entry)

XX Brain and Acute Leukemia, Cytoplasmic alternate protein #3.

XX acute myelogenous leukemia; gene expression; BALC;
XX chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
XX Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

XX

DR WPI; 2003-441564/41.
N-PSDB; ADM46953.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
PT in a patient comprising assaying for the overexpression of one or more
PT BALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 19; 78bp; English.

XX The invention relates to a method of characterizing acute myelogenous
CC leukemia (AML) in a patient by assaying for the overexpression of one or
CC more BALC transcripts in cells obtained from the patient, where an
CC overexpression indicates that the patient has an aggressive form of AML.
CC The methods, kits and probes are useful for characterizing acute or
CC chronic myelogenous leukemia, or prostate cancer. They are also useful
CC for detecting BALC overexpression. This sequence corresponds to a BALC
CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
CC spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 180 AA;

Query Match

80.5%; Score 617.5; DB 7; Length 180;

Best Local Similarity 67.2%; Pred. No. 1.7e-59; Mismatches 14; Indels 35; Gaps 1;

Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCGGSRADAIIPRYESWTRETSWLTYYTDSALPSAANTDSGPAGGLHAGVLEDP 53
|||||
1 MCGGSRADAIIPRYESWTRETSWLTYYTDSALPSAANTDSGPAGGLHAGVLEDP 60
Db 1 MCGGSRADAIIPRYESWTRETSWLTYYTDSALPSAANTDSGPAGGLHAGVLEDP 60
QY 54 -----GVLEDPSSNGVLRPAAPGSIAPPEKKNCGT 85
|||||
61 KIKAPTVSDVDEGLFSASMAPLAVFSGHMLDGLPSNGVPRSTAPGIPNPEKKNCGT 120
Db 61 KIKAPTVSDVDEGLFSASMAPLAVFSGHMLDGLPSNGVPRSTAPGIPNPEKKNCGT 120
QY 86 QCPNQSLSGSLTQKNGLTWTEAKRDKKMSAREVAISVTENIRQMDRSKRVTKNCIN 145
|||||
121 QCPNQSLSGSLTQKNGLTWTEAKRDKKMSAREVAISVTENIRQMDRSKRVTKNCIN 180
Db 121 QCPNQSLSGSLTQKNGLTWTEAKRDKKMSAREVAISVTENIRQMDRSKRVTKNCIN 180

RESULT 10

ADM46963 standard; protein; 149 AA.

XX ADM46963;

XX 03-JUN-2004 (first entry)

XX Brain and Acute Leukemia, Cytoplasmic alternate protein #5.

XX acute myelogenous leukemia; gene expression; BALC;
XX chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
XX Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

DR WPI; 2003-441564/41.

N-PSDB; ADM46955.

XX

PT	Characterizing acute or chronic myelogenous leukemia, or prostate cancer
PT	in a patient comprises assaying for the overexpression of one or more
PT	BALC transcripts in cells obtained from the patient.
PS	Diclosure; SEQ ID NO 21; 78bp; English.
XX	
CC	The invention relates to a method of characterizing acute myelogenous
CC	leukemia (AML) in a patient by assaying for the overexpression of one or
CC	more BALC transcripts in cells obtained from the patient, where an
CC	overexpression indicates that the patient has an aggressive form of AML.
CC	The methods, kits and probes are useful for characterizing acute or
CC	chronic myelogenous leukemia, or prostate cancer. They are also useful
CC	for detecting BALC overexpression. This sequence corresponds to a BALC
CC	(Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
CC	spliced RNA consisting of exons 1, 6 and 8.
XX	
SQ	Sequence 149 AA;
Query Match	61.7%; Score 473.5; DB 7; Length 149;
Best Local Similarity	66.0%; Pred. No. 1e-45;
Matches 95; Conservative 2; Mismatches 12; Indels 35; Gaps 1	
Qy	1 MGCGSRADATBPRYRESWTRTESTTWLTYSDDALPSAAATDSCPEAGLHA----- 53
Db	1 MGCGSRADATBPRYRESWTRTESTTWLTYSDDAPSPAAPPDSCEPAGLHSEAEKS 60
Qy	54 -----GVLEDPGSSNGVLRPAAPGGIAPNEKKNCGT 85
Db	61 KIKAPTDSVSDEGLFSASKMAPLVFSGMLEDGLPSNGVRSTAPGGIPNPKTKNCET 120
Qy	86 QCPNSQSILSSGPPLTQKONGLWTE 109
Db	121 QCPNPQSLSSGPLTQKONGLOTTE 144
RESULT 11	
ADQ81902	
ID	ADQ81902 standard; protein; 92 AA.
XX	
AC	ADQ81902;
XX	
DT	09-SEP-2004 (first entry)
XX	
DE	Human dioxygenase 10.12.
XX	
KM	Human; enzyme: dioxygenase 10.12; malignant tumour; inflammation;
KM	immunological disease; haemopathy; HIV infection.
OS	Homo sapiens.
XX	
PN	CN1344798-A.
XX	
PD	17-APR-2002.
XX	
PF	29-SEP-2000; 2000CN-00125495.
XX	
PR	29-SEP-2000; 2000CN-00125495.
XX	
PA	(SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
P1	Mao Y, Xie Y;
PI	
XX	
DR	WI; 2002-509506/55.
XX	
DR	N-PADB; ADQ81901.
XX	
XX	New polypeptide human dioxygenase 10.12 and polynucleotides encoding this
PT	polypeptide, useful for treating various diseases, such as malignant
PT	tumors, inflammations, immunological diseases, hemopathy and HIV
PT	infection.
XX	
PS	Claim 1; SEQ ID NO 2; 33pp; Chinese.
XX	
CC	The present invention discloses a new kind of polypeptide, human

CC	dioxygenase 10.12, polynucleotides encoding this polypeptide, a DNA
CC	recombination process to produce the polypeptide and antagonist against
CC	the polypeptide. The present invention also discloses the method of
CC	applying the polypeptide in treating various diseases, such as malignant
CC	tumours, inflammations, immunological diseases, haemopathy and HIV
CC	infection. The present sequence is the human dioxygenase 10.12.
CC	
XX	Sequence 92 AA;
XX	
XX	Query Match 43.7%; Score 335.5; DB 5; Length 92;
XX	Best Local Similarity 72.7%; Pred. No. 8.8e-29;
XX	Matches 64; Conservative 9; Mismatches 10; Indels 5; Gaps 1;
XX	
OY	58 DGPSSNGVLRLPAPAGGIAPNEKKNNCGTCCPNSSGLSGPLTKONGI,MTTEAKRDAKRM 117
DB	10 DGPETGN-----AVGGI,PNPEKKNCCERQCPNPSLSSGPLTKONGI,QTTEAKRDAKRM 64
OY	118 SAREVAISVTENIRQMDRSKRVTKNCIN 145
DB	65 PAKEVITINVTDSIQOMDRSRRTIKNCVN 92
XX	
XX	RESULT 12
ID	ADO48481 standard; protein; 54 AA.
AC	ADO48481;
XX	
XX	ADO48481;
XX	12-AUG-2004 (first entry)
DE	Mouse PTH responsive gene protein exon 2 splice variant.
XX	
XX	PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
KW	transgenic animal; osteopathic; gene therapy; osteoporosis; mouse;
KM	murine.
OS	Mus sp.
XX	
XX	MO2004044152-A2.
PN	27-MAY-2004.
XX	
XX	10-NOV-2003; 2003WO-US035655.
PP	
XX	12-NOV-2002; 2002US-042532P.
PR	
XX	(AMHP) WTEHT.
PA	
XX	Robinson JA, Stojanovic-Susulic V, Babij P, Murrills RJ;
PI	
XX	WPI: 2004-420299/39.
XX	DR N-PSDB; ADO48480.
DR	
XX	
XX	New nucleic acid fragment encoding a PAIGB polypeptide, useful in
PT	preparing a composition for diagnosing, treating or preventing bone
PT	related disorders, e.g., osteoporosis.
XX	
XX	Claim 9; SEQ ID NO 10; 163bp; English.
XX	
XX	The invention relates to a novel PTH responsive gene (PAIGB) fragment
CC	encoding a polypeptide. The invention further comprises: a chimeric
CC	construct comprising the isolated nucleic acid fragment operatively
CC	linked to suitable regulatory sequences; a host cell transformed with the
CC	chimeric construct; a vector comprising the nucleic acid fragment;
CC	obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC	fragment; an antibody that specifically binds to one or more epitopes of
CC	a PAIGB polypeptide; a composition for regulating bone-forming activity
CC	in a mammal comprising the nucleic acid fragment; polypeptide or antibody
CC	; an agent that alters the expression of PAIGB gene or polypeptide;
CC	determining whether an agent alters the expression of PAIGB mRNA;
CC	screening agents for effectiveness in altering expression of the nucleic
CC	acid fragment; screening for agents useful for treating bone related
CC	

disorders; evaluating the efficacy of a treatment of a bone related disorder in a subject; identifying polypeptides capable of binding to PAIGB; monitoring the effectiveness of treatment of a subject with a bone related agent; a transgenic animal comprising the DNA; an animal model for the study of bone density modulation comprising a first group of animals composed of the transgenic animal and a second group of control animals; studying bone mass determinants; studying the modulation of bone mass; studying an effect of PAIGB on bone disorders; identifying an agent for treating bone related disorders; identifying whether an agent which has bone forming activity; and a stably transfected cell line comprising two constructs, the first construct comprising a ligand binding domain linked to a DNA binding domain which is linked to an activation domain all of which expression is driven by a constitutive promoter, the second construct comprising multiple copies of DNA binding elements linked to a minimal promoter which is linked to PAIGB cDNA, where upon the addition of chemical inducer, transcription of PAIGB gene is induced. The PAIGB polynucleotide has osteoparathic activity. The PTH responsive gene may be used to treat disorders by gene therapy. The nucleic acid is useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis. This sequence represents a PTH responsive gene protein of the invention.

SQ Sequence 54 AA;

Query Match 38.3%; Score 294; DB 8; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.6e-24;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETESTWLTYSDDALPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIIPRYESWTRETESTWLTYSDDALPSAATDGPAGGLHAG 54

RESULT 13

ADM46962 standard; protein; 73 AA.

AC ADM46962;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #4.

XX acute myelogenous leukemia; gene expression; BAALC;

XX chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

XX Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PS (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

DR WPI, 2003-441564/41.

XX N-PSDB; ADM46954.

XX Characterizing acute or chronic myelogenous leukemia, or prostate cancer

XX in a patient comprises assaying for the overexpression of one or more

XX BAALC transcripts in cells obtained from the patient.

XX Disclosure; SEQ ID NO 20; 78pp; English.

XX The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful for detecting BAALC overexpression. This sequence corresponds to a BAALC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 73 AA;

Query Match 37.5%; Score 288; DB 7; Length 73;
Best Local Similarity 91.4%; Pred. No. 1.1e-23;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRETESTWLTYSDDALPSAATDGPAGGLHAGVLED 58
DB 1 MCGGSRADAIIPRYESWTRETESTWLTYSDDALPSAATDGPAGGLHAGVLED 58

RESULT 14

ADM46960 standard; protein; 54 AA.

AC ADM46960;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #2.

XX acute myelogenous leukemia; gene expression; BAALC;

XX chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

XX Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PS (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapell A;

DR WPI, 2003-441564/41.

XX N-PSDB; ADM46952.

XX Characterizing acute or chronic myelogenous leukemia, or prostate cancer

XX in a patient comprises assaying for the overexpression of one or more

XX BAALC transcripts in cells obtained from the patient.

XX Disclosure; SEQ ID NO 18; 78pp; English.

XX The invention relates to a method of characterizing acute myelogenous

XX leukemia (AML) in a patient by assaying for the overexpression of one or

XX more BAALC transcripts in cells obtained from the patient, where an

XX overexpression indicates that the patient has an aggressive form of AML.

XX The methods, kits and probes are useful for characterizing acute or

XX chronic myelogenous leukemia, or prostate cancer. They are also useful

XX for detecting BAALC overexpression. This sequence corresponds to a BAALC

XX (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative

XX spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 54 AA;

Query Match 36.2%; Score 278; DB 7; Length 54;

Best Local Similarity 94.4%; Pred. No. 9.2e-23;

Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGGSRADAIERYYESWTRETSTWLTYYTDSDALPSAAATDGPAGGLHAG 54

1 MGGGSRADAIERYYESWTRETSTWLTYYTDSDALPSAAATDGPAGGLHAG 54

Db 1 MGGGSRADAIERYYESWTRETSTWLTYYTDSDALPSAAATDGPAGGLHAG 54

RESULT 15

ADO48477

ID ADO48477 standard; protein; 54 AA.

XX ADO48477;

DT 12-AUG-2004 (first entry)

DE Human PTH responsive gene protein exon 2 splice variant.

XX PTH responsive gene; PAIGB, bone-forming; bone; bone density modulation;

KW transgenic animal; osteopathic; gene therapy; osteoporosis; human.

OS Homo sapiens.

XX MO2004044152-A2.

XX 27-MAY-2004.

XX 10-NOV-2003; 2003MO-US035655.

XX 12-NOV-2002; 2002US-0425532P.

XX (AMMP) WYETH.

XX Robinson JA, Stojanovic-Susulic V, Babi P, Murrills RJ;

XX WPI; 2004-420299/39.

XX N-PSDB; ADO48476.

XX New nucleic acid fragment encoding a PAIGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.

XX Claim 9; SEQ ID NO 6; 169pp; English.

XX The invention relates to a novel PTH responsive gene (PAIGB) fragment
CC encoding a polypeptide. The invention further comprises: a chimeric
CC construct comprising the isolated nucleic acid fragment operatively
CC linked to suitable regulatory sequences; a host cell transformed with the
CC chimeric construct; a vector comprising the nucleic acid fragment;
CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC obtaining a polypeptide; detecting the presence of the nucleic acid
CC fragment; an antibody that specifically binds to one or more epitopes of
CC a PAIGB polypeptide; a composition for regulating bone-forming activity
CC in a mammal comprising the nucleic acid fragment; polypeptide or antibody
CC; an agent that alters the expression of PAIGB gene or polypeptide;
CC determining whether an agent alters the expression of PAIGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain

CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.

SQ Sequence 54 AA;

Query Match 36.2%; Score 278; DB 8; Length 54;

Best Local Similarity 94.4%; Pred. No. 9.2e-23;

Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MGGGSRADAIERYYESWTRETSTWLTYYTDSDALPSAAATDGPAGGLHAG 54

1 MGGGSRADAIERYYESWTRETSTWLTYYTDSDALPSAAATDGPAGGLHAG 54

Db 1 MGGGSRADAIERYYESWTRETSTWLTYYTDSDALPSAAATDGPAGGLHAG 54

Search completed: April 12, 2005, 15:37:16
job time : 109.667 secs

This Page Blank (uspto)


```
/ LENGTH: 778
5198347-4
Query Match      12.1%; Score 92.5; DB 6; Length 778;
Best Local Similarity 33.0%; Pred. No. 0.12;
Matches 29; Conservative 12; Mismatches 44; Indels 3; Gaps 3;

QY 31 TSDALPSAATDSCPEAGGIAHGVLEDPSSNGV-LRPAAPGGINPEKKN-C-GTQCP 88
DB 352 TVSSDVPVSGKSGSTSSAHLAAGENGVEVHGTDEPEDEKADPKQDIEVKKQDT 411
QY 89 NSQSLSS-GPLTQKQKGLMTTEAKRDAK 115
DB 412 DRSQGLGPHTERATLGETHMEKDT 439

RESULT 3
5198347-4
/ Patent No. 5198347
/ APPLICANT: Miller, LOUIS H.;ADAMS, JOHN H.;KASLOW,
/ DAVIC C.;FANG, XIANGDOUG
/ TITLE OF INVENTION: DNA ENCODING PLASMIDIUM VIVAX AND
/ PLASMIDIUM KOWLESII DUFFY RECEPTOR
/ NUMBER OF SEQUENCES: 27
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/554,837
/ FILING DATE: 20-JUL-1990
/ SEQ ID NO:4
/ LENGTH: 778
5198347-4

Query Match      12.1%; Score 92.5; DB 6; Length 778;
Best Local Similarity 33.0%; Pred. No. 0.12;
Matches 29; Conservative 12; Mismatches 44; Indels 3; Gaps 3;

QY 31 TSDALPSAATDSCPEAGGIAHGVLEDPSSNGV-LRPAAPGGINPEKKN-C-GTQCP 88
DB 352 TVSSDVPVSGKSGSTSSAHLAAGENGVEVHGTDEPEDEKADPKQDIEVKKQDT 411
QY 89 NSQSLSS-GPLTQKQKGLMTTEAKRDAK 115
DB 412 DRSQGLGPHTERATLGETHMEKDT 439

RESULT 4
US-09-328-352-8115
/ Sequence 8115, Application US/09328352
/ Patent No. 6562958
/ GENERAL INFORMATION:
/ APPLICANT: Gary L. Breton et al.
/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
/ TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS
/ FILE REFERENCE: GTC99-03PA
/ CURRENT APPLICATION NUMBER: US/09/328,352
/ CURRENT FILING DATE: 1999-06-04
/ NUMBER OF SEQ ID NOS: 8252
/ SEQ ID NO 8115
/ LENGTH: 726
/ TYPE: PRT
/ ORGANISM: Acinetobacter baumannii
US-09-328-352-8115

Query Match      11.0%; Score 84; DB 4; Length 726;
Best Local Similarity 35.1%; Pred. No. 1;
Matches 27; Conservative 8; Mismatches 22; Indels 20; Gaps 4;

QY 2 GCGSRADAIEPRYVSWTRETSTWLTYSDDALPSA--AATDSG-----DEAGGIAHG 54
DB 181 GFGAGGEDVWEPNDNVNMGDEK--WLAHRNSALAGSNLAATTEMGLIYVNE----- 231
QY 55 VLEDGSSNGVLRPAAP 71
DB 232 ----GPOAGDPRSAAP 244
```

```
RESULT 5
US-09-408-647A-2
/ Sequence 2, Application US/09408647A
/ Patent No. 6399858
/ GENERAL INFORMATION:
/ APPLICANT: Kobayashi, Donald
/ TITLE OF INVENTION: Chitinase Gene from Stenotrophomonas
/ TITLE OF INVENTION: maltophilia
/ FILE REFERENCE: Rut-Cook 98-0090
/ CURRENT APPLICATION NUMBER: US/09/408,647A
/ CURRENT FILING DATE: 1999-08-26
/ PRIOR APPLICATION NUMBER: 60/098,036
/ PRIOR FILING DATE: 1998-08-27
/ NUMBER OF SEQ ID NOS: 2
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 2
/ LENGTH: 700
/ TYPE: PRT
/ ORGANISM: Stenotrophomonas maltophilia
/ FEATURE:
/ NAME/KEY: SIGNAL
/ LOCATION: (1)...(41)
/ NAME/KEY: DOMAIN
/ LOCATION: (196)...(290)
/ NAME/KEY: DOMAIN
/ LOCATION: (330)...(483)
/ PUBLICATION INFORMATION:
/ DATABASE ACCESSION NUMBER: Genbank No. 6399858 AF014950
/ DATABASE ENTRY DATE: 1997-09-23
US-09-408-647A-2

Query Match      10.9%; Score 83.5; DB 3; Length 700;
Best Local Similarity 24.0%; Pred. No. 1.1;
Matches 31; Conservative 14; Mismatches 57; Indels 27; Gaps 2;

QY 26 TWLTYSDDALPSAATDSCPEAGGIAHGVLEDPSSNGVLRPAAPGGINPEKKN-C--- 81
DB 159 TMANASAGSHTKAVATDNNNNAVTSATVSVTVTASNDTTPSPVGGIASSKRTATTVN 218
QY 82 -----NCG-----TCCPNSQSLSSGPLTQKQKGLMTTEAKRDAKMS 118
DB 219 LVWSAATDNGSGGVAGYDVYRNGSLVGSFATQYTDGLTFATAYTVTRARDNAGNAS 278
QY 119 AREVAISVT 127
DB 279 AQSGSISVT 287

RESULT 6
US-09-071-035-176
/ Sequence 176, Application US/09071035
/ Patent No. 6448043
/ GENERAL INFORMATION:
/ APPLICANT: Gil H. Choi
/ TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
/ NUMBER OF SEQUENCES: 496
/ CORRESPONDENCE ADDRESS:
/ ADDRESSER: Human Genome Sciences, Inc.
/ STREET: 9410 Key West Avenue
/ CITY: Rockville
/ STATE: Maryland
/ COUNTRY: USA
/ ZIP: 20850
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
/ COMPUTER: HP Vectra 486/33
/ OPERATING SYSTEM: MSDOS version 6.2
/ SOFTWARE: ASCII Text
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/09/071,035
/ FILING DATE:
```


CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 176:
SEQUENCE CHARACTERISTICS:
LENGTH: 286 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-176

Query Match 10.4%; Score 79.5; DB 4; Length 286;
Best Local Similarity 26.6%; Pred. No. 0.89; Indels 41; Gaps 9;
Matches 47; Conservative 16; Mismatches 73; Indels 41; Gaps 9;

QY 3 CGG-----SRADAIEPRYYESW---TRETSTWLTYSDDLPSAAAT-DSGPEAG 49
DB 1 CGGKSTENTDSSSAESTTVESTKASATKSSSK-ATTKSDAKPSGTTTADSKATAS 59
QY 50 GLHAGVLEDPSSNGVLRPAAPGIANPEKKN-CGTQCPN---SGLSSGPTLQKNGU 105
DB 60 STKEA-----ANNGSAEKQSPAKNANPDQANOVNLQANMPGQGLPQALITSTNNF 113
QY 106 WTEAKRDARKMSAREVAISVTENIRQMD-RSKRYT-----KNCIN 145
DB 114 LTAATTSQADQNNFRLVYAEKEALPVNDARVNQLTPISSFEKTYGSDAEAKNAVN 170

RESULT 7
US-09-071-035-174
Sequence 174, Application US/09071035
GENERAL INFORMATION:
APPLICANT: Gil H. Choi
TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
NUMBER OF SEQUENCES: 496
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4MB storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/071,035
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: A. Anders Brookes
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB369P2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 174:
SEQUENCE CHARACTERISTICS:

LENGTH: 305 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-174

Query Match 10.4%; Score 79.5; DB 4; Length 305;
Best Local Similarity 26.6%; Pred. No. 0.98; Indels 41; Gaps 9;
Matches 47; Conservative 16; Mismatches 73; Indels 41; Gaps 9;

QY 3 CGG-----SRADAIEPRYYESW---TRETSTWLTYSDDLPSAAAT-DSGPEAG 49
DB 20 CGGKSTENTDSSSAESTTVESTKASATKSSSK-ATTKSDAKPSGTTTADSKATAS 78
QY 50 GLHAGVLEDPSSNGVLRPAAPGIANPEKKN-CGTQCPN---SGLSSGPTLQKNGU 105
DB 79 STKEA-----ANNGSAEKQSPAKNANPDQANOVNLQANMPGQGLPQALITSTNNF 132
QY 106 WTEAKRDARKMSAREVAISVTENIRQMD-RSKRYT-----KNCIN 145
DB 133 LTAATTSQADQNNFRLVYAEKEALPVNDARVNQLTPISSFEKTYGSDAEAKNAVN 189

RESULT 8
US-08-374-077C-2
Sequence 2, Application US/08374077C
Patent No. 6027912
GENERAL INFORMATION:
APPLICANT: Hall, Linda M.
APPLICANT: Ren, Dejian
APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha1
TITLE OF INVENTION: Calcium Channel Subunit
NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/374,077C
FILING DATE: 19-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-374-077C-2

Query Match 10.2%; Score 78.5; DB 3; Length 2516;
Best Local Similarity 22.4%; Pred. No. 27;
Matches 39; Conservative 21; Mismatches 61; Indels 53; Gaps 8;
QY 3 CGGSRADAIEPRY--YESWTRETSTWLTYSDDLPSAAATDSGPEAGGLHAGVLEDP 60

Db 207 CCGGGISAPPRLTPEAWOLOPQ-----NSVTSAGSTNSPSSSGG-----GR 249
Qy 61 SSGVLRPAAPGSIAPPEKKMC-----GTQC-----PNS 90
Db 250 DNNSSY--SAVGDDSSSNCSNCDITGDNSTLHGLGVGVCSFIADCDNSEDDGDPNN 307
Qy 91 QSLSSGPL-TOKONGMTTEAKRDARMSAREVAISVTENIRQMDRSKRVTKNC 143
Db 308 QDLSSQTLRTAIVAAVAALAAKEQAQESLADCE-SFSDRQDADDEVRITIDQC 360

RESULT 9

US-08-895-590-2
; Sequence 2, Application US/08895590
; Patent No. 6207410
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
; NUMBER OF SEQUENCES: 101
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/895,590
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,888
; FILING DATE: 19-JAN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 022650-263
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-895-590-2

Query Match 10.2%; Score 78.5; DB 3; Length 2516;
Best Local Similarity 22.4%; Pred. No. 27;
Matches 39; Conservative 21; Mismatches 61; Indels 53; Gaps 8;

Qy 3 CCGSRADAIEPRY--YESWTRTESTTWLTYSDALPSAAATDSGPEAGLHAGVLEDP 60
Db 207 CCGGGISAPPRLTPEAWOLOPQ-----NSVTSAGSTNSPSSSGG-----GR 249
Qy 61 SSGVLRPAAPGSIAPPEKKMC-----GTQC-----PNS 90
Db 250 DNNSSY--SAVGDDSSSNCSNCDITGDNSTLHGLGVGVCSFIADCDNSEDDGDPNN 307
Qy 91 QSLSSGPL-TOKONGMTTEAKRDARMSAREVAISVTENIRQMDRSKRVTKNC 143
Db 308 QDLSSQTLRTAIVAAVAALAAKEQAQESLADCE-SFSDRQDADDEVRITIDQC 360

RESULT 10
US-09-539-879A-2
; Sequence 2, Application US/09539879A
; Patent No. 6436627
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Calcium Channel Subunit

NUMBER OF SEQUENCES: 57
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,879A
FILING DATE: 31-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/111,865
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/374,077
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-539-879A-2

Query Match 10.2%; Score 78.5; DB 4; Length 2516;
Best Local Similarity 22.4%; Pred. No. 27;
Matches 39; Conservative 21; Mismatches 61; Indels 53; Gaps 8;

Qy 3 CCGSRADAIEPRY--YESWTRTESTTWLTYSDALPSAAATDSGPEAGLHAGVLEDP 60
Db 207 CCGGGISAPPRLTPEAWOLOPQ-----NSVTSAGSTNSPSSSGG-----GR 249
Qy 61 SSGVLRPAAPGSIAPPEKKMC-----GTQC-----PNS 90
Db 250 DNNSSY--SAVGDDSSSNCSNCDITGDNSTLHGLGVGVCSFIADCDNSEDDGDPNN 307
Qy 91 QSLSSGPL-TOKONGMTTEAKRDARMSAREVAISVTENIRQMDRSKRVTKNC 143
Db 308 QDLSSQTLRTAIVAAVAALAAKEQAQESLADCE-SFSDRQDADDEVRITIDQC 360

RESULT 11
US-09-121-964-1
; Sequence 1, Application US/09121964
; Patent No. 6124447
; GENERAL INFORMATION:
; APPLICANT: Natori, Shunji
; TITLE OF INVENTION: NOVEL ENZYME CATALYZING DEPHOSPHORYLATION

```

; FILE REFERENCE: 32290-144753
; CURRENT APPLICATION NUMBER: US/09/121,964
; CURRENT FILING DATE: 1998-07-24
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 724
; TYPE: PRT
; ORGANISM: Sarcophaga peregrina
US-09-121-964-1

```

```

Query Match          10.2%; Score 78; DB 3; Length 724;
Best Local Similarity 28.6%; Pred. No. 5.1;
Matches 33; Conservative 6; Mismatches 40; Indels 34; Gaps 4;

```

```

QY 25 STWLTITDSDALPSAATDSCP-EAGGLHAGVLEDDG-----PSSNGV 65
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2 STTLEVTQKDLHDLLEIFKKPFDGSCERKSLSEEDITNRCFLFKKDYTLIELDNGV 61
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 66 LRPAAPGSIANPE-----KMNCGTCPCNSQSLSSGPTLTKONGL 105
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 62 LSPRYPERIRITPEYEHGHTSTTPNNTNLGTQATNHAPOSG--KNGNGL 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 12
US-09-914-259-10
; Sequence 10, Application US/09914259
; Patent No. 6495336
; GENERAL INFORMATION:
; APPLICANT: Makowski, Lee
; APPLICANT: Hyman, Paul
; APPLICANT: Williams, Mark
; TITLE OF INVENTION: STAGED ASSEMBLY OF NANOSTRUCTURES
; FILE REFERENCE: 8471-010-999
; CURRENT APPLICATION NUMBER: US/09/914,259
; CURRENT FILING DATE: 2000-11-21
; NUMBER OF SEQ ID NOS: 180
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10
; LENGTH: 885
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-914-259-10

```

```

Query Match          10.2%; Score 78; DB 4; Length 885;
Best Local Similarity 29.6%; Pred. No. 6.8;
Matches 34; Conservative 16; Mismatches 33; Indels 32; Gaps 7;

```

```

QY 49 GGHAGVLEDDGPSNGVLRPAAPGI-----ANPEKMNCGTCPCNSQSLSSGP---L 98
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 381 GSITHS---DKPT--ILRPATVGGTLEDDGTOAKQKAPVSE--SQSAGAPMAA 431
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 99 TOKONGIWTTEAKR-----DAKMSAREVAISVTEN---IRQMDRSEKVT 140
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 432 TQCKEGYSPSRKRGPLSKLMAEDGERTSARAVLTIVKDDHGLDQFSSNVAVS 486
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 13
US-09-220-528-104
; Sequence 104, Application US/09220528A
; Patent No. 6284540
; GENERAL INFORMATION:
; APPLICANT: Milbrandt, Jeffrey D.
; APPLICANT: Baloh, Robert H.
; TITLE OF INVENTION: Artemin, A No. 6284540e1 Neurotrophic Factor
; FILE REFERENCE: 6029-7998
; CURRENT APPLICATION NUMBER: US/09/220,528A
; CURRENT FILING DATE: 1998-12-24
; EARLIER APPLICATION NUMBER: 09/218,698
; EARLIER FILING DATE: 1998-12-22
; EARLIER APPLICATION NUMBER: 60/108,148
; EARLIER FILING DATE: 1998-11-12
; EARLIER APPLICATION NUMBER: 09/163,283

```

```

; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 120
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 104
; LENGTH: 215
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-220-528-104

```

```

Query Match          10.1%; Score 77.5; DB 3; Length 215;
Best Local Similarity 26.5%; Pred. No. 1;
Matches 31; Conservative 12; Mismatches 41; Indels 33; Gaps 5;

```

```

QY 5 GSRADAIER-----YIESWTRETESTWLTITDSDALPSAATDSCPFRAGLHA----- 53
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 65 GADAAALGGRGCAACARSWCRCASAWATAPTSVCVASAAPAARA--LHTTSAMP 122
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 54 GVLDEDPSSNGVLRPAAPGSIANPEKMNCGTCPCNSQSLSSGPTLTKONGLWTTA 110
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 123 AYVAPPCD---RPPAPGPSASP-----AADPRATKSPSWTSTA 159
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 14
US-09-620-405B-473
; Sequence 473, Application US/09620405B
; Patent No. 6528054
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; FILE REFERENCE: 210121.470C8
; CURRENT APPLICATION NUMBER: US/09/620,405B
; CURRENT FILING DATE: 2000-07-20
; NUMBER OF SEQ ID NOS: 495
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-620-405B-473

```

```

Query Match          10.1%; Score 77.5; DB 4; Length 445;
Best Local Similarity 21.5%; Pred. No. 2.9;
Matches 37; Conservative 28; Mismatches 58; Indels 49; Gaps 7;

```

```

QY 7 RADAIER-----RYIESWTRETESTWLTITDSD-ALPSAATDSCPFRAGLHAGVLEDDG 60
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 78 RADEILPSBSKQDYESSWDSLSCEIVSQKOVCLPKAHQKEIDKIN---GKLEESP 133
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 61 SSGVLRPAAPGSIANPEKMNCGTCPCNSQSLSSGPTLTKONGLWTTA-----NCGTCPCPS 90
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 134 DNDGFLKAPCRMKVSIPTALBLMDNQTFRKAEPPKPSAFEPALIEKQKVPNALVLEK 193
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 91 QSLSSG---PLTKONGL-----WTEAKRDAKMSAREVAISVTENIRQMDR 135
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 194 QTLRADQMPSPSKQKQKVENSWDSLSLET--VSGKDVCPVPAATKQKEMDK 243
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

RESULT 15
US-09-433-826B-473
; Sequence 473, Application US/09433826B
; Patent No. 6579973
; GENERAL INFORMATION:
; APPLICANT: Jiang, Yuguu
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Xu, Jiaqichun
; APPLICANT: Harlocker, Susan L.

```

```

; TITLE OF INVENTION: COMPOSITIONS FOR THE TREATMENT AND
; TITLE OF INVENTION: DIAGNOSIS OF BREAST CANCER AND METHODS FOR THEIR USE
; FILE REFERENCE: 210121.4700C4
; CURRENT APPLICATION NUMBER: US/09/433,826B
; CURRENT FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 474
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 473
; LENGTH: 445
; TYPE: PR1
; ORGANISM: Homo sapiens
; US-09-433-826B-473

```

Query Match	10.1%;	Score 77.5;	DB 4;	Length 445;
Best Local Similarity	21.5%;	Pred. No. 2.9;		
Matches 37;	Conservative 28;	Mismatches 58;	Indels 49;	Gaps 7;

```

0Y      , RAAAIIE-----RYYSSWIRRETESRWLJYTBSD -ALPSAAATBDSGPFGAGLHNGVLDEGP 60
Db      78 RAEIILPSESKQKQDYESSMSDESLCETVSOQKDYCLPRAHQKEIKDIN-----GKLEESP 133
0Y      61 SSGVLRPAAPGAIAPKKMM-----NCGTQCPNS 90
Db      134 DINGFLTAPCKMKVSLPTVALTLMQITFKAPPEKPSAEPALTEMQSVNNALIELKNE 193
0Y      91 QSLSSG--PLTQKQNGL-----WTEAKRDAKAMREVALSTENTENROMDR 135
Db      194 QTRAPAOQMPSESKQKQKVEENMSDESLRET--VSQKQVCYPAKTHQEQENDK 243

```

Search completed: April 12, 2005, 15:43:28
Job time : 27.6667 secs

Job time : 27.66667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:37:30 ; Search time 75.6667 Seconds
(without alignments)
636.206 Million cell updates/sec

Title: US-10-705-716A-2

Perfect score: 767
Sequence: 1 MCGGSRADAIERYESWT.....VTENIRQMDRSKVTKNKCN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:
1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	645	84.1	145	US-10-293-239-17	Sequence 17, Appl
2	645	84.1	145	US-10-177-390-30	Sequence 30, Appl
3	617.5	80.5	160	US-10-293-239-19	Sequence 19, Appl
4	473.5	61.7	149	US-10-293-239-21	Sequence 21, Appl
5	288	37.5	73	US-10-293-239-20	Sequence 20, Appl
6	278	36.2	54	US-10-293-239-18	Sequence 18, Appl
7	272	35.5	80	US-10-293-239-18	Sequence 22, Appl
8	140	18.3	25	US-10-293-239-37	Sequence 37, Appl
9	99	12.9	18	US-10-293-239-35	Sequence 35, Appl
10	94	12.3	147	US-10-767-701-48905	Sequence 48905, A
11	85.5	11.1	274	US-10-425-114-65588	Sequence 65588, A
12	85	11.1	450	US-10-437-963-137075	Sequence 137075, A
13	84	11.0	718	US-10-282-122A-45268	Sequence 45268, A

14	83	10.8	1433	9	US-09-801-368-60	Sequence 60, Appl
15	83	10.8	1433	15	US-10-149-310-112	Sequence 112, Appl
16	82.5	10.8	539	15	US-10-424-599-227399	Sequence 227399, Appl
17	82.5	10.8	609	16	US-10-437-963-144983	Sequence 144983, Appl
18	82	10.7	1491	16	US-10-437-963-140197	Sequence 140197, Appl
19	82	10.7	1597	16	US-10-437-963-180225	Sequence 180225, Appl
20	81.5	10.6	600	16	US-10-408-765A-1945	Sequence 1945, Appl
21	81	10.6	670	14	US-10-156-76A-14107	Sequence 14107, Appl
22	80.5	10.5	320	15	US-10-425-114-43318	Sequence 43318, A
23	80.5	10.5	508	16	US-10-437-963-188141	Sequence 188141, A
24	79.5	10.4	286	9	US-09-071-035-176	Sequence 176, Appl
25	79.5	10.4	286	14	US-10-206-576-176	Sequence 176, Appl
26	79.5	10.4	286	17	US-10-912-362-117	Sequence 174, Appl
27	79.5	10.4	305	9	US-09-071-035-174	Sequence 174, Appl
28	79.5	10.4	305	14	US-10-206-576-174	Sequence 174, Appl
29	79.5	10.4	305	17	US-10-912-362-117	Sequence 174, Appl
30	79.5	10.4	661	13	US-10-007-805-552	Sequence 552, Appl
31	79.5	10.4	661	14	US-10-076-622-552	Sequence 552, Appl
32	79.5	10.4	661	14	US-10-124-805-552	Sequence 552, Appl
33	79.5	10.4	1013	13	US-10-007-805-553	Sequence 553, Appl
34	79.5	10.4	1013	14	US-10-076-622-553	Sequence 553, Appl
35	79.5	10.4	1013	14	US-10-124-805-553	Sequence 553, Appl
36	78.5	10.2	262	16	US-10-424-599-242710	Sequence 242710, Appl
37	78	10.2	355	15	US-10-424-599-242710	Sequence 242710, Appl
38	78	10.2	455	15	US-10-087-192-1317	Sequence 1317, Appl
39	78	10.2	661	16	US-10-451-467A-554	Sequence 554, Appl
40	78	10.2	832	13	US-10-087-192-1317	Sequence 1317, Appl
41	78	10.2	885	14	US-10-080-608A-11	Sequence 10, Appl
42	78	10.2	1489	15	US-10-370-685-99	Sequence 99, Appl
43	78	10.2	1479	16	US-10-437-963-180217	Sequence 180217, Appl
44	77.5	10.1	215	9	US-09-320-920-104	Sequence 104, Appl
45	77.5	10.1	445	9	US-09-825-301-29	Sequence 29, Appl

ALIGNMENTS

RESULT 1	US-10-293-239-17	US-10-293-239-17
1	Sequence 17, Application US/10293239	
2	Publication No. US20030119043A1	
3	GENERAL INFORMATION:	
4	APPLICANT: Tanner, Stephan	
5	APPLICANT: de la Chapelle, Albert	
6	TITLE OF INVENTION: BAALC expression as a diagnostic marker for acute leukemia	
7	FILE REFERENCE: 22727/04101	
8	CURRENT APPLICATION NUMBER: US/10/293,239	
9	CURRENT FILING DATE: 2002-11-12	
10	PRIOR APPLICATION NUMBER: US 60/346,210	
11	PRIOR FILING DATE: 2001-11-09	
12	NUMBER OF SEQ ID NOS: 39	
13	SOFTWARE: PatentIn version 3.1	
14	SEQ ID NO 17	
15	LENGTH: 145	
16	TYPE: PRT	
17	ORGANISM: Homo sapiens	
18	US-10-293-239-17	
19	Query Match	84.1%; Score 645; DB 14; Length 145;
20	Best Local Similarity	83.4%; Pred. No. 7, 1e-57;
21	Matches	121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
22	QY	1 MCGGSRADAIERYESWTRETSTWLTYSDDALPSAATDSGEAGLHAGVLEDP 60
23	DB	1 MCGGSRADAIERYESWTRETSTWLTYSDDALPSAATDSGEAGLHAGVLEDP 60
24	QY	61 SSGVGRAPAPGGIINAPKKMGCGTCCPSQSISSEPLTQKNGLTTRAKRPAKXMSAR 120
25	DB	61 PSNGVGRSTRPAGGIEPKEKKNCECTCPNPQSSSPPLTQKNGLTTRAKRPAKXMPAK 120
26	QY	121 EVAISVTENIRQMDRSKVTKNKCN 145
27	DB	121 EYTIINTDSIQMDRSRRTTKNKN 145

RESULT 2

US-10-177-390-30
; Sequence 30, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerold
; APPLICANT: N.V. Antwerp Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with linear
; FILE REFERENCE: 021505wc/JH/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-30

Query Match 84.1%; Score 645; DB 14; Length 145;
Best Local Similarity 83.4%; Pred. No. 7.1e-57;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 60
DB 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 60
QY 61 SSNGVSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 120
DB 61 PSNGVSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 120
QY 121 EVAISTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCIN 145

RESULT 3

US-10-293-239-19
; Sequence 19, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 19
; LENGTH: 180
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-19

Query Match 80.5%; Score 617.5; DB 14; Length 180;
Best Local Similarity 67.2%; Pred. No. 5.5e-54;
Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 53
DB 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 60
QY 54 -----GVLBDGSSNGVLRPAAPGAIANPEKKNCGT 85
DB 61 KIKAPTDVSDGLFSASKAPLAIVSHQMLDGLPSNGVPRSTAPGGIIPNEKKNCGT 120
QY 86 QCPNQSLSGFLTKQKNGILQTTAKRDAKRMASREVAISVTENIRQMDRSKRVTKNCIN 145

DB 121 QCPNQSLSGFLTKQKNGILQTTAKRDAKRMASREVAISVTENIRQMDRSRRITKNCIN 180

RESULT 4

US-10-293-239-21
; Sequence 21, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-21

Query Match 61.7%; Score 473.5; DB 14; Length 149;
Best Local Similarity 60.0%; Pred. No. 1.4e-39;
Matches 95; Conservative 2; Mismatches 12; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 53
DB 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 60
QY 54 -----GVLBDGSSNGVLRPAAPGAIANPEKKNCGT 85
DB 61 KIKAPTDVSDGLFSASKAPLAIVSHQMLDGLPSNGVPRSTAPGGIIPNEKKNCGT 120
QY 86 QCPNQSLSGFLTKQKNGILQTTAKRDAKRMASREVAISVTENIRQMDRSKRVTKNCIN 145
DB 121 QCPNQSLSGFLTKQKNGILQTTAKRDAKRMASREVAISVTENIRQMDRSRRITKNCIN 180

RESULT 5

US-10-293-239-20
; Sequence 20, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 73
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-20

Query Match 37.5%; Score 288; DB 14; Length 73;
Best Local Similarity 91.4%; Pred. No. 2.6e-21;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 58
DB 1 MCGGSRADAIERPRYESTWRETSTWLTYSDDLPSAAATDGPAGGLHAGVLEDP 58

RESULT 6

```
US-10-293-239-18
; Sequence 18, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-18

Query Match          36.2%; Score 278; DB 14; Length 54;
Best Local Similarity 94.4%; Pred. No. 1,8e-20;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGGRADAIEPRYESTWRETSTWLTGTTSDALPSAATDGPAGGLHAG 54
Db 1 MCGGGRADAIEPRYESTWRETSTWLTGTTSDAPPSAAPDGPAGGLHAG 54

RESULT 7
US-10-293-239-22
; Sequence 22, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-22

Query Match          35.5%; Score 272; DB 14; Length 80;
Best Local Similarity 94.3%; Pred. No. 1,2e-19;
Matches 50; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGGRADAIEPRYESTWRETSTWLTGTTSDALPSAATDGPAGGLHAG 53
Db 1 MCGGGRADAIEPRYESTWRETSTWLTGTTSDAPPSAAPDGPAGGLHAG 53

RESULT 8
US-10-293-239-37
; Sequence 37, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; CURRENT FILING DATE: 2002-11-12
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-37

Query Match          12.9%; Score 99; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0049;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 9 DAIEPRYESTWRETST 26
Db 1 DAIEPRYESTWRETST 18

RESULT 10
US-10-767-701-48905
; Sequence 48905, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovacic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21 (53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 48905
; LENGTH: 147
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; OTHER INFORMATION: Clone ID: LIR3476-020-P1-K1-A12 pep
US-10-767-701-48905

Query Match          12.3%; Score 94; DB 16; Length 147;
Best Local Similarity 27.5%; Pred. No. 0.22;
Matches 30; Conservative 17; Mismatches 52; Indels 10; Gaps 3;

QY 36 LPSAATDSGRAGGAGVLEDPSSNGVLRPAAPGIANREKKNCTGCPNQS----
```

Db 29 LPAALPEDAGEAGADNGAGHGPARG-----QABAGAGARGALHCGPCPHSDSCIG 84
QY 93 --LSSGFLQKQ-NGMTTEAKRDARKMARREVAISVTENIRQMDSKR 138
Db 85 RMTSTGTSGHRPCSGLLSVQLRQMPRSSAEQVREKIKQIDQORRVR 133

RESULT 11

US-10-425-114-65588
; Sequence 65588, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 65588
; LENGTH: 274
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4863-011-Fl12_Fli1.pep
US-10-425-114-65588

Query Match 11.1%; Score 85.5; DB 15; Length 274;
Best Local Similarity 32.4%; Pred. No. 3.4;
Matches 35; Conservative 10; Mismatches 42; Indels 21; Gaps 5;

QY 32 DSDALPSAATDSCGPAAGLHAGVLEBDGSSNGVLRPAAPGCI---ANPEKKNCGTQCP 88
Db 147 DSDDAAPAAAHHDDGPAVVAAGLGGSSSG-LPPAAGAATAAEFSLSLGLPLP 205

QY 89 -----NSQISSGFLQKQNGMTTEAKRDARKMARREV 122
Db 206 AAEPAEAADDESRRNQGAAS-PLIEEGEG--NAQLAVRRVRREV 250

RESULT 12

US-10-437-963-137075
; Sequence 137075, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 137075
; LENGTH: 450
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38593C.1.pep
US-10-437-963-137075

Query Match 11.1%; Score 85; DB 16; Length 450;
Best Local Similarity 25.6%; Pred. No. 7.1;
Matches 45; Conservative 14; Mismatches 53; Indels 64; Gaps 8;

QY 4 GGSRADAI-----EPRIYESTRTREESTWLTYYTSDALPSAAATDSCGPAAGLH 52
Db 3 GGS-ADAVTKEMBALVGCNPVNAVSGETCETSSKKGKVAADSGSHSSPPEDDDAOG-- 59

QY 53 AGVLEDGSSNGVLRPAAPGCIANPEKKNCGTQCPNSQSLSSGFLTO----- 100
Db 60 -----DGSQDG-----GSEPAKKKKK-----SKSKKKGFLQOTDPPSITIDELF 101

QY 101 -----KONGLM--TTEAKRDARKN-----SAREVAISVTENIRQMDRS 136
Db 102 PGSDPPEGEIQGYKODNLMRTTSEKRELERLQKPMYNAVRAAEVHRQVRGRMS 157

RESULT 13

US-10-282-122A-45268
; Sequence 45268, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zykkind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Foreyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR FILING DATE: 2000-05-23
; PRIOR FILING DATE: 2000-07-27
; PRIOR FILING DATE: 2000-05-26
; PRIOR FILING DATE: 2000-03-35
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 45268
; LENGTH: 718
; TYPE: PRT
; ORGANISM: Acinetobacter baumannii
US-10-282-122A-45268

Query Match 11.0%; Score 84; DB 15; Length 718;
Best Local Similarity 35.1%; Pred. No. 16;
Matches 27; Conservative 8; Mismatches 22; Indels 20; Gaps 4;

QY 2 GCGSGRADAIPEPRYESTRTREESTWLTYYTSDALPSA--AATDSC-----PEAGGLHAG 54

Db 173 GFGAGEDVWEPDNDVWNGDEKE--WLAHNSBALAGSNLAATEMGLIYVNP----- 223
QY 55 VLEDGSSNGVLRPAAP 71
Db 224 ----GFGAGSDPRSAAP 236

RESULT 14
US-09-801-368-60
Sequence 60, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Bueby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587
PRIOR FILING DATE: 1999-10-20
NUMBER OF SEQ ID NOS: 440
SOFTWARE: Patentin version 3.0
SEQ ID NO 60
LENGTH: 1433
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-09-801-368-60

Query Match 10.8%; Score 83; DB 9; Length 1433;
Best Local Similarity 29.4%; Pred. No. 49;
Matches 32; Conservative 15; Mismatches 50; Indels 12; Gaps 4;
QY 36 LPSAATDSGPAGGLHAGVLEDPSSNGVLRPAAGGIANPEKKNCGTQCPSQSLS 95
Db 927 LPATTSLKPLFGSGSKSLNRQRTPNVK-----ENPEHEYLGNDSNNNNSEA 979
QY 96 G--PLTQKONGLWTTAKDKAKMSAREVAISTENIR--QMDRSKRYT 140
Db 980 GSPMTNTTNGNKRKLKYEKDAKR-NAKDGISKGENAHNFQNDTKKMS 1027

RESULT 15
US-10-149-310-112
Sequence 112, Application US/10149310
Publication No. US20040077039A1
GENERAL INFORMATION:
APPLICANT: Holtzman, Douglas
APPLICANT: Madden, Kevin T.
APPLICANT: Maxon, Mary
APPLICANT: Sherman, Amir
TITLE OF INVENTION: Modulation of Secondary Metabolite Production by
FILE REFERENCE: 14184-019US1
CURRENT APPLICATION NUMBER: US/10/149,310
CURRENT FILING DATE: 2003-02-19
PRIOR APPLICATION NUMBER: PCT/US01/29288
PRIOR FILING DATE: 2001-09-19
PRIOR APPLICATION NUMBER: US 60/233,564
PRIOR FILING DATE: 2000-09-19
NUMBER OF SEQ ID NOS: 308

SOFTWARE: Patentin version 3.1
SEQ ID NO 112
LENGTH: 1433
TYPE: PRT
ORGANISM: Saccharomyces cerevisiae
US-10-149-310-112

Query Match 10.8%; Score 83; DB 15; Length 1433;
Best Local Similarity 29.4%; Pred. No. 49;
Matches 32; Conservative 15; Mismatches 50; Indels 12; Gaps 4;
QY 36 LPSAATDSGPAGGLHAGVLEDPSSNGVLRPAAGGIANPEKKNCGTQCPSQSLS 95
Db 927 LPATTSLKPLFGSGSKSLNRQRTPNVK-----ENPEHEYLGNDSNNNNSEA 979
QY 96 G--PLTQKONGLWTTAKDKAKMSAREVAISTENIR--QMDRSKRYT 140
Db 980 GSPMTNTTNGNKRKLKYEKDAKR-NAKDGISKGENAHNFQNDTKKMS 1027

Search completed: April 12, 2005, 16:03:42
Job time : 79.6667 secs

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:28:59 ; Search time 21.6667 Seconds
(without alignments)
643.912 Million cell updates/sec

Title: US-10-705-716A-2

Perfect score: 767

Sequence: 1 MCGSGRADAIAPRYESWT.....VTENIRQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92.5	12.1	571	2 T43456	hypothetical prote
2	92.5	12.1	778	2 A35970	erythrocyte-bindin
3	88	11.5	365	1 GNV6SC	genome polyprotein
4	83	10.8	1433	2 S54587	CAR8 protein - yea
5	81	10.6	269	2 T37073	hypothetical prote
6	81	10.6	967	2 S66852	hypothetical prote
7	79.5	10.4	2559	2 T09144	probable guanine n
8	78.5	10.2	1159	2 A40670	nuclear envelope p
9	78	10.2	885	2 T09225	A kinase anchor pr
10	78	10.2	3488	2 T34418	hypothetical prote
11	76.5	10.0	600	2 C69899	conserved hypochet
12	76	9.9	645	2 S19156	serotonin receptor
13	76	9.9	839	2 B84624	hypothetical prote
14	75.5	9.8	499	2 S22571	integrase-like pro
15	75.5	9.8	1063	2 T03743	biological protein -
16	75.5	9.8	1122	2 T47424	hypothetical prote
17	75	9.8	368	1 TVMSGML	transforming prote
18	75	9.8	521	2 T51693	XPolycarb - Africa
19	74.5	9.7	601	2 AH0784	probable transport
20	74.5	9.7	4957	2 T03455	ALR protein - huma
21	74.5	9.7	5262	2 T03454	ALR protein - huma
22	74	9.6	260	2 B38594	crotonin I - fruit
23	74	9.6	742	2 T38001	probable phosphat
24	74	9.6	832	2 T31878	hypothetical prote
25	74	9.6	962	2 T00262	hypothetical prote
26	73.5	9.6	876	1 A57988	regulatory protein
27	73.5	9.6	960	2 T37916	probable heterochr
28	73.5	9.6	1575	2 S68448	synaptojanin, 170k
29	73	9.5	403	2 H98327	enantiomer-selecti

30	73	9.5	453	2 AE2955	glutamy] tRNA amid
31	73	9.5	781	2 T41551	hypothetical prote
32	73	9.5	1274	2 T37193	enamelin matrix pr
33	72.5	9.5	251	2 C75521	cytochrome-related
34	72.5	9.5	284	2 T51172	transcription fact
35	72.5	9.5	435	2 AG1028	prepillin imported
36	72.5	9.5	503	2 T35053	probable solute-bi
37	72.5	9.5	516	2 JEO301	inulinase (EC 3.2.
38	72.5	9.5	1343	2 AP0611	cell division prot
39	72.5	9.5	3623	2 T08618	intrinsc factor-B
40	72.5	9.5	5327	2 T13564	microtubule-associ
41	72	9.4	108	2 T31565	hypothetical prote
42	72	9.4	1038	1 JC5757	DNA-directed DNA p
43	72	9.4	1038	2 T18222	DNA polymerase del
44	72	9.4	1573	2 T50113	3-dehydroquinate e
45	71.5	9.3	381	2 AB3048	8-amino-7-oxononan

ALIGNMENTS

RESULT 1

T43456
hypothetical protein DKFZp434L061.1 - human

C:Species: Homo sapiens (man)

C>Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #ext_change 09-Jul-2004

C:Accession: T43456

R:Pouetka, A.; Klein, M.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.

submitted to the Protein Sequence Database, December 1999

A:Reference number: 222516

A:Accession: T43456

A>Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-571 <AAA>

A:Cross-references: UNIPROT:075175; EMBL:AL133647

A:Experimental source: adult testis; clone DKFZp434L061

C:Genetics:

A>Note: DKFZp434L061.1

C:Superfamily: hydroxyproline-rich glycoprotein

Query Match

Best Local Similarity 12.1%; Score 92.5; DB 2; Length 571;

Matches 30; Conservative 7; Mismatches 32; Indels 25; Gaps 3;

QY	31	TSQDALPSAAATDSCPEAGGLH-----AGVLEDPSSNGVTRPAP--	71
DB	111	TSSEVSQSPAKNGSKVHSHQHPSPAVPPTYPSPGPPAASALSTTPGNNGVAPAPPS	170
QY	72	--GGTANPEKMNCGTQCPNSQSLG----SGPLT	99
DB	171	ALGPKASPAASHNSGTAPAYAQVAPAPSGEST	204

RESULT 2

A35970
erythrocyte-binding protein - Plasmodium knowlesi

C:Species: Plasmodium knowlesi

C>Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #ext_change 09-Jul-2004

C:Accession: A35970

R:Adams, J.H.; Hudson, D.E.; Torii, M.; Ward, G.E.; Wellems, T.E.; Aikawa, M.; Miller, T.

Cell 63, 141-153, 1990

A>Title: The duffy receptor family of plasmodium knowlesi is located within the micronem

A:Reference number: A35970; MIDID:91004213; PMID:2170017

A:Accession: A35970

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-778 <ADA>

A:Cross-references: UNIPROT:P22545; GB:M68518; GB:M37513; NID:9160273; PID:9160274

C:Keywords: transmembrane protein

Query Match

Best Local Similarity 12.1%; Score 92.5; DB 2; Length 778;

Matches 29; Conservative 12; Mismatches 44; Indels 3; Gaps 3;

A:Residues: 1-849,862-885 <D02>
 A:Cross-references: EMBL:AF033275; NID:g2852698; PID:g2852699
 A:Note: binds the regulatory subunit (RI) of protein kinase AII isoforms
 A:Accession: T09227
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: mRNA
 A:Residues: 1-789, 'PGGHTG' <D03>
 A:Cross-references: EMBL:AF033276; NID:g2852700; PID:g2852701
 A:Note: binds the regulatory subunit (RI) of protein kinase AII isoforms
 C:Genetics:
 A:Gene: AKAP-XL
 C:Keyword: alternative splicing; kidney; lung; signal transduction

Query Match 10.2%; Score 78; DB 2; Length 885;
 Best Local Similarity 29.6%; Pred. No. 34;
 Matches 34; Conservative 16; Mismatches 33; Indels 32; Gaps 7;

QY 49 GGIHAGVLEDPSSNGVLRPAAGT-----ANPEKKNCGTQCPNSQSLSSGP--L 98
 DB 381 GSIHS-----DKPPT--ILRPATVGGTLEDGQTQAAKEQKAPCVSE--SOSAGAGPANA 431

QY 99 TOKONGMTTEAR-----DAKMSAREVAISVTEN---IKOMRSKRYT 140
 DB 432 TQKEGPPYSEPSKRGPLSKLMAEDGERTSARAVLTWVKDHDGILDPSSRSVNV 486

RESULT 10
 T34418
 hypothetical protein F12F3.3 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 29-Oct-1999
 C:Accession: T34418
 R:Fulton, B.; Wohldmann, P.
 Submitted to the EMBL Data Library, July 1998
 A:Description: The sequence of C. elegans cosmid F12F3.
 A:Reference number: Z21521
 A:Accession: T34418
 A:Status: preliminary; translated from GB/EMBL/DBD
 A:Molecule type: DNA
 A:Residues: 1-3488 <F01>
 A:Cross-references: EMBL:U00022; PID:AAC25885.1; GSPDB:GN00023; CESP:F12F3.3
 A:Experimental source: strain Bristol N2; clone F12F3
 C:Genetics:
 A:Gene: CESP:F12F3.3
 A:Map position: 5
 A:Introns: 281/3; 332/1; 562/3; 600/3; 1866/3; 1944/3; 3393/1

Query Match 10.2%; Score 78; DB 2; Length 3488;
 Best Local Similarity 24.3%; Pred. No. 1.6e+02;
 Matches 33; Conservative 23; Mismatches 36; Indels 44; Gaps 7;

QY 24 ESTWL--TYTDSALPSAATDSGP-----EAGGLHAGVLEDPSSNGVLRPAAGCIAN 76
 DB 2772 DSDMLEIANTDRKKFKDRSLTBSGEVYVQVATGTAH-----VSSPESEETN 2817

QY 77 PEKKKNCGTQCPNSQSLSSGPTLQKONGMTTEAKDARKMSAREVAI-----S 125
 DB 2818 PVKILVPGSEMPASK-----TEKK-----TDPAKSESGKAEIYAEKQDQGSAS 2866

QY 126 VTENIRQMRSKRYTK 141
 DB 2867 TTEAVEE-KTKKVKV 2881

RESULT 11
 C69899
 conserved hypothetical protein yobL - Bacillus subtilis
 C:Species: Bacillus subtilis
 C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
 C:Accession: C69899
 R:Kunet, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter, C.; Bron, S.; Broutelle, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho, A.; Ehrlich, S.D.; Emmerson, P.T.; Ertian, K.D.; Erington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997
 A:Authors: Foulger, D.; Filtz, C.; Fujita, M.; Fujita, Y.; Funa, S.; Gallizzi, A.; Gallier, J.; Harwood, C.R.; Hentut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Huilo, M.P.; Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kunita, K.; Lapidus, A.; Lardinois, A.; Muthers, Lander, U.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Maeda, S.; Mauei, Y.; M.; Ogawa, K.; Ogawa, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle, Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon, A.; Authors: Schleich, S.; Schotter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Sexor, A.; Authors: Schleich, S.; Schotter, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Sexor, T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasunoto, K.; Yata, K.; Yoshida, K.
 A:Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
 A>Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
 A:Reference number: A69580; MUID:198044033; PMID:9384377
 A:Accession: C69899
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown
 A:Molecule type: DNA
 A:Residues: 1-600 <KUN>
 A:Cross-references: UNIPROT:Q34330; GB:Z99114; GB:AL009126; NID:g2634230; PID:CA13792.
 A:Experimental source: strain 168
 C:Genetics:
 A:Gene: yobL

Query Match 10.0%; Score 76.5; DB 2; Length 600;
 Best Local Similarity 23.1%; Pred. No. 30;
 Matches 39; Conservative 23; Mismatches 56; Indels 51; Gaps 7;

QY 9 DAIEPRYESTRE-----TESTWLTYSALPSAATDSGPEAGLHAGVLEDPSS 61
 DB 342 DLSAIEBSYQKMDVNGDAYSRMYT-----ALGSVAVAVVGRK-----AGALNKADA 393

QY 62 SNGVLRPAAGGIA-----NPEKKNCGTQCP-----NSQSLSSGPT----- 99
 DB 394 AGKVINKASQAKKIKDKVILPDLPPNPKYKALADVPPNVVDSQNLNELLTNAKKIP 453

QY 100 -----QKONGMTTEAKDARKMSAREVAISVTENIRQMRSKRY 139
 DB 454 DGRKPFQKKSPMLNKEYD-----AYELGRKAKKGVKDVSRV 497

RESULT 12
 S19156
 serotonin receptor 2B - fruit fly (Drosophila melanogaster)
 N:Alternate names: 5-hydroxytryptamine receptor 2B (5-HT2B)
 C:Species: Drosophila melanogaster
 C>Date: 07-Apr-1994 #sequence_revision 07-Apr-1994 #text_change 09-Jul-2004
 C:Accession: S19156; S18154
 R:Saudou, F.; Bouchet, U.; Amlaiky, N.; Plaseat, J.L.; Hen, R.
 EMBO J. 11, 7-17, 1992
 A>Title: A family of Drosophila serotonin receptors with distinct intracellular signal
 A:Reference number: S19156; MUID:92155185; PMID:1310937
 A:Accession: S19156
 A:Molecule type: mRNA
 A:Residues: 1-645 <SAT>
 A:Cross-references: UNIPROT:P28286; EMBL:Z11490; NID:g7506; PID:CAA77571.1; PID:g7507
 C:Genetics:
 A:Gene: FLYBase:5-HT1B
 A:Cross-references: FLYBase:FBgn0004572
 A:Superfamily: octopamine receptor type I
 C:Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
 F:124-145/Domain: transmembrane #status predicted <TM1>
 F:156-177/Domain: transmembrane #status predicted <TM2>
 F:193-214/Domain: transmembrane #status predicted <TM3>
 F:234-256/Domain: transmembrane #status predicted <TM4>
 F:284-305/Domain: transmembrane #status predicted <TM5>
 F:564-587/Domain: transmembrane #status predicted <TM6>
 F:597-619/Domain: transmembrane #status predicted <TM7>

Query Match 9.9%; Score 76; DB 2; Length 645;
 Best Local Similarity 31.0%; Pred. No. 36;
 Matches 31; Conservative 12; Mismatches 31; Indels 26; Gaps 4;

QY 29 TYTDSALPSAATDSGPEAGLHAGVLEDPSSNGVLRPAAG-----IANPEKKNCG 84

DB 489 TTTPSEKALSGAGTVAGAVAGSGSGSEEGAGTEGKAGVGLASTIANPHOKL--- 545
QY 85 TQCPNQSGLSSGPLTQKQNGLMTTEAKRKDAKMSAREVAI 124
DB 546 -----AKRQL--LEAKRE--RRAAGTLAI 566

RESULT 13

E84824
hypothetical protein At2g40040 [imported] - Arabidopsis thaliana
C/Spectes: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C/Accession: E84824
R/Lin: X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Bentol, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Ko, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.;
eues, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.
Nature 402, 761-768, 1999
A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A/Reference number: A84420; MUID:20083487; PMID:10617197
A/Accession: E84824

A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-839 <STO>
A/Cross-references: UNIPROT:O04207; GB:AB002093; NID:g2088657; PIDN:AAB95289.1; GSPDB:GN
C/Genetics:
A/Gene: At2g40040
A/Map position: 2

Query Match 9.4%; Score 76; DB 2; Length 839;
Best Local Similarity 23.8%; Pred. No. 49;
Matches 40; Conservative 14; Mismatches 62; Indels 52; Gaps 6;

QY 5 GSRADAIERYRYSWTR---ETESTWLTITDSDALPSAAA----- 41
DB 338 GSGAGVGLP-----WKKSESTESNGATWGSXGAAMNSWDKNIETDSEPAWG 392
QY 42 -----TDSGEAGGL---HAGVLEDGSSNGV-----LRPAAPGGIANPEKKM 81
DB 393 SQCKNSSETSGRAAMGAMWKKSESTEPGAGMGMDKNSSETELGPAAMGMDKKSDT 452
QY 82 NCQTQCPNQSGLSSGPLTQKQNGLMTTEAKRKDAKMSAREVAISVTEN 129
DB 453 KSGPAAMWSTDAAMAGSSDKN---SETSDAAMGSRNKKTSIES 496

RESULT 14

S22571
integrase-like protein PE65 - rat
C/Spectes: Rattus norvegicus (Norway rat)
C/Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C/Accession: S22571; S22572
R/Duilio, A.; Zambreno, N.; Mogavero, A.R.; Ammendola, R.; Cimino, F.; Russo, T.
Nucleic Acids Res. 19, 5269-5274, 1991
A/Title: A rat brain mRNA encoding a transcriptional activator homologous to the DNA birt
A/Reference number: S22571; MUID:9200215; PMID:1923810
A/Accession: S22571

A/Molecule type: mRNA
A/Residues: 1-499 <DU11>
A/Cross-references: UNIPROT:Q99MK3; EMBL:X60469; NID:g57559; PIDN:CAA42999.1; PID:g57560
A/Accession: S22572
A/Status: nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 31-318 <DU12>
A/Cross-references: EMBL:X60468; NID:g57561; PIDN:CAA42998.1; PID:g1177617
A/Note: this sequence was submitted to the EMBL Data Library, July 1991
C/Genetics:
A/Intons: 88/3; 107/3; 135/3; 157/3; 207/3; 250/2; 252/2; 290/3
C/Keywords: transcription regulation
P/42-78/Domain: WW repeat homology <WW1>

Query Match 9.8%; Score 75.5; DB 2; Length 499;
Best Local Similarity 25.6%; Pred. No. 30;
Matches 34; Conservative 17; Mismatches 59; Indels 23; Gaps 5;

QY 13 PRYSEWTESTETWLTITDSDALPSAAATDSCPEAGGLHAGVLEDGSSNGVLRPAAPG 72
DB 77 PQGNSPOEESQLTITGTF-----AHQGFEEGER-----WDEPEEAPMEL----- 118

QY 73 GIAPBEKKKNCGTQCPNQSGLSSGPLTQKQNGLMTTEAKRKDAKMSAREVA-ISVTENIR 131
DB 119 GLKDPPE-----CTLFPASQSLSPFVPOEEENLPQRNANPGIKCFVAVSLGVEMTEEBEL 174

QY 132 QMDRSKRVTKNCI 144
DB 175 AGRSSSVAVANNCI 187

RESULT 15

T03743
bifocal protein - fruit fly (Drosophila melanogaster)
C/Spectes: Drosophila melanogaster
C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 09-Jul-2004
C/Accession: T03743
R/Bahr, S.M.; Yang, X.Y.; Chia, W.
Mol. Cell. Biol. 17, 5521-5529, 1997
A/Title: The Drosophila bifocal gene encodes a novel protein which colocalizes with acti
A/Reference number: Z15048; MUID:97415628; PMID:9271427
A/Accession: T03743

A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1063 <BAH>
A/Cross-references: UNIPROT:O16125; EMBL:AF011793; NID:g2388667; PIDN:AAB69991.1; PID:g2
C/Genetics:
A/Note: bifocal
A/Gene: FlyBase:FBgn0014133

Query Match 9.8%; Score 75.5; DB 2; Length 1063;
Best Local Similarity 20.7%; Pred. No. 71;
Matches 39; Conservative 23; Mismatches 59; Indels 67; Gaps 6;

QY 5 GSRADAIERYRYSWTR---ETESTWLTITDSDAL-----PSAAAT- 42
DB 83 GAIDTFEPATISTSQKRNMTGSEBKSEKSIISNTSDSTGCHHSVAVASLPDAAATT 142
QY 43 -----DSGPEAGGLHAGVLEDGSSNGVLRPAAPG 72
DB 143 NYTVPIPKQSSSLNTRSGEREMRYILSESGERDGLSEGEQPAQGVVSNRCGEVETG 202
QY 73 GIAPBEKKKNCGTQCPNQSGLSSGPLTQKQNGLMTTEAKRKDAKMSAREVAISVTENIRQ 132
DB 203 TIGSPSSSAN---QNPNPWHLK---TKCKPGQSVAEGRPSAKE-----TIVNSKS 247
QY 133 MDRSKRV 140
DB 248 CSKTKSIS 255

Search completed: April 12, 2005, 15:44:43
Job time : 26.6667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:16:23 ; Search time 94.6667 Seconds

(without alignments)
784.347 Million cell updates/sec

Title: US-10-705-716a-2

Perfect score: 767

Sequence: 1 MCGGSGRADAIERYRYESWT.....VTENIKRMDRSKRVTKNCIN 145

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot 03:*

1: uniprot_sprot:*

2: uniprot_tramb1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	767	100.0	145	Q920K5	Q920K5 rattus norv
2	746	97.3	145	Q8VH1	Q8VH1 mus musculu
3	645	84.1	145	Q9HA33	Q9HA33 homo sapien
4	634	82.7	145	Q8WNE9	Q8WNE9 sus scrofa
5	617.5	80.5	180	Q8WXS3	Q8WXS3 homo sapien
6	473.5	61.7	149	Q8WXS1	Q8WXS1 homo sapien
7	459	59.8	123	Q9CY89	Q9CY89 mus musculu
8	294	38.3	54	Q8VBS8	Q8VBS8 mus musculu
9	294	38.3	54	Q790N3	Q790N3 rattus norv
10	288	37.5	73	Q8WXS0	Q8WXS0 homo sapien
11	278	36.2	54	Q8WTP6	Q8WTP6 homo sapien
12	272	35.5	80	Q8WXS2	Q8WXS2 homo sapien
13	267	34.8	54	Q8WNE8	Q8WNE8 sus scrofa
14	161	21.0	32	Q801V5	Q801V5 brachydantio
15	92.5	12.1	613	Q6ZM46	Q6ZM46 homo sapien
16	92.5	12.1	753	CNO3_HUMAN	CNO3_HUMAN
17	92.5	12.1	1073	PVDA_PLANK	PVDA_PLANK
18	88	11.5	1035	POLG_SUNVS	POLG_SUNVS
19	87	11.3	1035	Q76C74	Q76C74 sugarcane m
20	87	11.3	1713	Q8TGE1	Q8TGE1 saccharomyc
21	86	11.2	585	Q7U2Z9	Q7U2Z9 rhodospirillum rubrum
22	86	11.2	1070	PVDG_PLANK	PVDG_PLANK
23	84.5	11.0	534	Q96SA2	Q96SA2 homo sapien
24	84	11.0	155	Q6USF5	Q6USF5 plasmodium falciparum
25	83.5	10.9	214	Q8WTC2	Q8WTC2 leucophaea
26	83.5	10.9	677	Q6A015	Q6A015 mus musculu
27	83.5	10.9	700	Q30678	Q30678 xanthomonas
28	83.5	10.9	751	CNO3_MOUSE	CNO3_MOUSE
29	83	10.8	581	Q89EL1	Q89EL1 bradyrhizobium
30	83	10.8	581	P89204	P89204 sugarcane m
31	83	10.8	1433	CAT8_YEAST	CAT8_YEAST

32	82.5	10.8	514	2	Q7SEF2	Q7SEF2 neurospora
33	82.5	10.8	853	1	AOP2_MOUSE	AOP2_MOUSE mus musculu
34	82.5	10.8	1537	2	Q6KXFO	Q6KXFO debrayomyce
35	82	10.7	339	2	Q8KRB5	Q8KRB5 streptomyces
36	82	10.7	882	1	AREA_ASPNG	AREA_ASPNG
37	82	10.7	1449	2	Q7XPB1	Q7XPB1 aspergillus
38	81.5	10.6	397	2	Q94KA7	Q94KA7 phaseolus v
39	81.5	10.6	852	1	AOP2_HUMAN	AOP2_HUMAN
40	81.5	10.6	888	2	Q8W0U4	Q8W0U4 borghum bic
41	81.5	10.6	3064	2	Q8B388	Q8B388 sugarcane m
42	81.5	10.6	3105	2	Q8U219	Q8U219 cercopithec
43	81	10.6	228	2	Q91TQ3	Q91TQ3 tupaiid her
44	81	10.6	269	2	Q951Y3	Q951Y3 streptomyces
45	81	10.6	670	2	Q828T2	Q828T2 streptomyces

ALIGNMENTS

RESULT 1	Q920K5	PRELIMINARY;	PRT;	145 AA.
AC	Q920K5			
DT	01-DEC-2001 (TEMBREL. 19, Created)			
DT	01-DEC-2001 (TEMBREL. 19, Last sequence update)			
DT	05-JUL-2004 (TEMBREL. 27, Last annotation update)			
DE	Dam-A20-4 (BALC Isoform 1-6-8).			
GN	Name=dem-A20-4; Synonyms=BALC;			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	(1)			
RP	SEQUENCE FROM N.A.			
RA	Wang X., Tian Q., Li W., Okano A., Suzuki T.;			
RL	Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=Sprague-Dawley;			
RC	MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;			
RA	Tanner S.M., Austin J.L., Leone G., Rush J.E., Plasse C., Heinonen K.,			
RA	Manozer K., Sill H., Kautela S., Kolitz J.E., Archer K.J.,			
RA	Califuri M.A., Bloomfield C.D., de la Chapelle A.;			
RT	"BALC, the human member of a novel mammalian neurotrophin gene			
RT	lineage, is implicated in hematopoiesis and acute leukemia.";			
RL	Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).			
DR	EMBL; AB073318; BAB70507.1; -			
DR	EMBL; AF371321; AAL50517.1; -			
DR	PFam; PF06989; BALC_N.1.			
SQ	SEQUENCE 145 AA; 15475 MW; DSA27AD67456F341 CRC64;			
Query Match	100.0%; Score 767; DB 2; Length 145;			
Best Local Similarity	100.0%; Pred. No. 2.8e-61;			
Matches	145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	1 MCGGSGRADAIERYRYESWTRETSWLTYSDDALPSAATSGPAGGLHGVLEDP 60			
DB	1 MCGGSGRADAIERYRYESWTRETSWLTYSDDALPSAATSGPAGGLHGVLEDP 60			
QY	61 SSNGVLRPAAPGGIAPDEKMGCTCPNSQSLSGFLTKQNGLTTEAKRDKRSAR 120			
DB	61 SSNGVLRPAAPGGIAPDEKMGCTCPNSQSLSGFLTKQNGLTTEAKRDKRSAR 120			
QY	121 EVAISVTENIKRMDRSKRVTKNCIN 145			
DB	121 EVAISVTENIKRMDRSKRVTKNCIN 145			
RESULT 2	Q8VH1	PRELIMINARY;	PRT;	145 AA.
ID	Q8VH1			
AC	Q8VH1			
DT	01-MAR-2002 (TEMBREL. 20, Created)			

DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DB 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE BAAIC isoform 1-6-8 (Mus musculus 16 days neonate cerebellum cDNA,
RIKEN full-length enriched library, clone:963002H16 product:brain and
DN acute leukemia, cytoplasmic, full insert sequence).
GN Name=BaalC
OS Mus musculus (Mouse)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21574583; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
RA Callisuri M.A., Bloomfield C.D., de la Chapelle A.,
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=9279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.,
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=2049374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komno H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:11617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Komno H., Akiyama J., Nishi K., Kitsuami T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watabiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.,
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:11757-1771(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Komno H., Kouda M., Koya S.,

RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ono M., Ohnaka N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai K., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.,
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF371320; AL50516.1; -
DR EMBL; AK079337; BAC37611.1; -
DR MGI; 1928704; BaalC.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 15515 MW; 4972670A618C4D6D CRC64;
Query Match 97.3%; Score 746; DB 2; Length 145;
Best local Similarity 97.2%; Pred. No. 2,2e-59;
Matches 141; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
DB 1 MCGGSRADAIIPRYESWTRETESTWLTYSDDLPSAAATDSGPAGLHAGVLEDP 60
1 MCGGSRADAIIPRYESWTRETESTWLTYSDDLPSAAATDSGPAGLHAGVLEDP 60
QY 61 SSGVLRPAAPGIANPERKNNCGTCPPSSQSLSSGPELTQKNGIMTTEAKRDKMSAR 120
DB 61 SSGVLRPAAPGIANPERKNNCGTCPPSSQSLSSGPELTQKNGIMTTEAKRDKMSAR 120
QY 121 EVAISYENIRQMDRSKRYTKNCIN 145
DB 121 EVAISYENIRQMDRSKRYTKNCIN 145
RESULT 3
Q9HA93 PRELIMINARY; PRT; 145 AA.
AC Q9HA93;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein FL12015 (BAAIC isoform 1-6-8) (Brain and acute
GN leukemia, cytoplasmic) (BAAIC 1-6-8).
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole embryo;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohtsuka M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
RA Niimoriya K., Ishihashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hirano K., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yorida M., Houta T., Kusano J.,
RA Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsumawa H., Ichihara T., Shiohara N., Sano S.,
RA Moriya S., Montiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiro M., Omori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsunaga K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togsashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,

```

RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuo Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isega T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs."
RL Nat. Genet. 36:40-45(2004).
RN [2]
RN SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnae.241525498;
RX Tanner S.M., Austen J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
RX Mroczek K., Sill H., Knutti S., Koltz J.E., Archer K.J.;
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALBC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia."
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [3]
RN SEQUENCE FROM N.A.
RN TISSUE=Brain;
RX MEDLINE=22288557; PubMed=12477932; DOI=10.1073/pnae.242603899;
RX Strunberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,
RA Altchul S.F., Zeeberg B., Buero K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marinina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldi M.F., Cabaavatt T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Plange C.,
RA Rahn S.S., Loggellano N.A., Peters G.J., Abramson R.D., Millaby S.J.,
RA Bosak S.A., McEwen P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S.S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakeley A.C., Touchman J.W., Green E.D., Dickson M.C.;
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smallos D.E., Scherch A., Schein J.E.,
RA Jones S.J., Maira M.A.;
RA "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RN SEQUENCE FROM N.A.
RN TISSUE=Brain;
RA Strunberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK022077; BAB13960.1; -
DR EMBL; AF371319; AAL50515.1; -
DR EMBL; BC011517; AAH1517.1; -
DR EMBL; AF365378; AAL50377.1; -
DR InterPro; IPR009728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 15551 MW; CFB92BBE283D92E CRC64;

Query Match 84.1%; Score 645; DB 2; Length 145;
Best Local Similarity 83.4%; Pred. No. 2.7e-50;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0

QY 1 MCGGSGRADATIPRYESWTRETSWTLYTSDALPSAAATDSGPEAGLJAGVLEDP 60
DB 1 MCGGSGRADATIPRYESWTRETSWTLYTSDAPSPSAAADPSGEAGLHSGMLEDL 60
QY 61 SSGVGLRPAAPGGINPEKKNACGTQCPNQSGLSSGPTLKQKNGLTTEAKSKDAKMSAR 120
DB 61 PBNQVPRSTAPRGINPEKKNCTCFQCPRPQSLSSGPTLKQKNGLTTEAKSKDAKMPAK 120
QY 121 EVAISVTENIRQMDRSKRYTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 4
Q8WNE9 PRELIMINARY; PRT; 145 AA.
Q8WNE9;

```

DT	01-MAR-2002 (Tremblrel. 20, Created)
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE	BALC isoform 1-6-8.
GN	Name=BALC;
OS	Sus scrofa (Pig).
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
OX	NCBI_TaxID=9823;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21574584; Pubmed=11707601; DOI=10.1073/pnas.241525498;
RA	Tanner S.M., Austen J.L., Leone G., Rugh L.J., Plase C., Heinonen K.,
RA	Mozek K., Still H., Knuttila S., Koiltz J.E., Archer K.O.,
RA	Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT	"BALC," the human member of a novel mammalian neuroectoderm gene
RT	lineage, is implicated in hematopoiesis and acute leukemia.";
RL	Proc. Natl. Acad. Sci. U.S.A. 98:113901-113906(2001).
DR	EMBL; AF311332; AAU50158.1;-
DR	Pfam; PF06989; BALC.N.1.
SQ	SEQUENCE 145 AA; I5401 MW; C7SEDTD00EPE2E26 CRC64;
<hr/>	
Query Match	82.7%; Score 634; DB 2; Length 145;
Best Local Similarity	82.8%; Pred. No. 2,6e-49;
Matches 120; Conservative 10; Mismatches 15; Indels 0; Gaps 0;	
OY	1 MGCGSGRADAIERPYYESWTRTESTSTWLTYTDSAPLPSAANTDSCPEAGGLHACVLEDGP 60
DB	1 MGCGSGRADAIERPYYESWTRTESTSTWLTYTDSAPPSPNAAPDSGPEAGGLQAQAVLDGV 60
OY	61 SSNGVLPPAAGGIANEKKNKCCCTGCPCNSGSGLGTOKONGLTMTFAKRDKRMSAR 120 :::
DB	61 SANGPSPSTAGGTSNPBKMKSCCTQCPCNPPOSLSGPLQKNGLRITTEARRDKAKRSIAK 120 :::
OY	121 EVAISVTENIRQMDSRKVKTKNCIN 145 :::
DB	121 EVTINVTESIRIQVDNRNQRTKKCN 145 :::
<hr/>	
RESULT 5	
OBWXS3	PRELIMINARY; PRT; 180 AA.
AC	OBWXS3;
ID	OBWXS3;
DT	01-MAR-2002 (Tremblrel. 20, Created)
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE	BALC 1-5-6-8.
GN	Name=BALC;
OS	Homo sapiens (Human).
OC	Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX	NCBI_TaxID=9606;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=21574584; Pubmed=11707601; DOI=10.1073/pnas.241525498;
RA	Tanner S.M., Austen J.L., Leone G., Rugh L.J., Plase C., Heinonen K.,
RA	Mozek K., Still H., Knuttila S., Koiltz J.E., Archer K.O.,
RA	Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT	"BALC," the human member of a novel mammalian neuroectoderm gene
RT	lineage, is implicated in hematopoiesis and acute leukemia.";
RL	Proc. Natl. Acad. Sci. U.S.A. 98:113901-113906(2001).
DR	EMBL; AF313578; AAU50379.1;-
DR	Pfam; PF06989; BALC.N.1.
SQ	SEQUENCE 180 AA; I9224 MW; 380193E0F186F684 CRC64;
<hr/>	
Query Match	80.5%; Score 617.5; DB 2; Length 180;
Best Local Similarity	67.2%; Pred. No. 1e-47;
Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;	
OY	1 MGCGSGRADAIERPYYESWTRTESTSTWLTYTDSAPLPSAANTDSCPEAGGLHA----- 53
DB	1 MGCGSGRADAIERPYYSWTRTESTSTWLTYTDSAPPSPAAPDSGPEAGGLHSVLEAEKS 60

```
QY 54 -----GVLEDPSSNGVLRPAAPGGIANPEKXNCGT 85
DB 61 KIKAPDVSDEGLFSAKMAPLAVFSGWLEDPSSNGVLRPAAPGGIANPEKXNCGT 120
QY 86 QCPNSQSLSSGPTOKONGMTTEAKRDAKMSAREVAISVTENINQMDRSKRVTKNCIN 145
DB 121 QCPNPQSLSSGPTOKONGMTTEAKRDAKMSAREVAISVTENINQMDRSKRVTKNCIN 180

RESULT 6
Q8KXSI PRELIMINARY; PRT; 149 AA.
AC Q8KXSI;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DB BAALC 1-5-6-7-8.
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
(1)
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50381.1; -.
DR Genev; HGNC:14333; BAALC.
PFam; PF06989; BAALC.N.1.
SQ SEQUENCE 149 AA; 15577 MW; DB26C1969B9B1F6E CRC64;

Query Match 61.7%; Score 473.5; DB 2; Length 149;
Best Local Similarity 66.0%; Pred. No. 7.5e-35;
Matches 95; Conservative 2; Mismatches 12; Indels 35; Gaps 1;

QY 1 MCCGGSRAAIPRRYVESWTRETSTWLTYSDDALPSAANDSGEAGGLA----- 53
DB 1 MCCGGSRAAIPRRYVESWTRETSTWLTYSDDALPSAANDSGEAGGLSVLEAKS 60
QY 54 -----GVLEDPSSNGVLRPAAPGGIANPEKXNCGT 85
DB 61 KIKAPDVSDEGLFSAKMAPLAVFSGWLEDPSSNGVLRPAAPGGIANPEKXNCGT 120
QY 86 QCPNSQSLSSGPTOKONGMTTE 109
DB 121 QCPNPQSLSSGPTOKONGMTTE 144

RESULT 7
Q9CYS9 PRELIMINARY; PRT; 123 AA.
AC Q9CYS9;
DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)
DB Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DB enriched library, clone:2810457D07 product:brain and acute leukemia,
DB cytoplasmic, full insert sequence.
GN Name=Baalc;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
(1)
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
```

```
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RA The FANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20459374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.,
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN (5)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Alzawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuana T., Teshiro H., Itoh M.,
RA Sumi N., Iehi Y., Nakamura S., Hazama M., Nishine T., Harada Y.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN (6)
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RA Adachi J., Alzawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hangaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013358; BAB28808.1; -.
DR MGD; MGI:1928704; Baalc.
SQ SEQUENCE 123 AA; 13391 MW; 33DFCB5F8CFAA0AB CRC64;

Query Match 59.8%; Score 459; DB 2; Length 123;
Best Local Similarity 94.6%; Pred. No. 1.2e-33;
Matches 88; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 53 AGVLEDPSSNGVLRPAAPGGIANPEKXNCGTQCPNSQSLSSGPTOKONGMTTEAKR 112
DB 31 SGVLEDPSSNGVLRPAAPGGIANPEKXNCGTQCPNSQSLSSGPTOKONGMTTEAKR 90
QY 113 DAKMSAREVAISVTENINQMDRSKRVTKNCIN 145
DB 91 DAKMSAREVAISVTENINQMDRSKRVTKNCIN 123

RESULT 8
```

Q8VBS8
ID Q8VBS8 PRELIMINARY; PRT; 54 AA.
AC Q8VBS8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAAIC isoform 1-8.
GN Name=BAAIC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=129S6/SvEvTac;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
R Mozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371324; AAL50520.1; -.
DR Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR MGI; MGI:1928704; Baaic.
DR Pfam; PF06989; BAAIC N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BFBCDB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.3e-19;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESWTRETESTLTYTDSALPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIEPRYESWTRETESTLTYTDSALPSAATDGPAGGLHAG 54

RESULT 9
Q790N3
ID Q790N3 PRELIMINARY; PRT; 54 AA.
AC Q790N3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BAAIC isoform 1-8.
GN Name=BAAIC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
R Mozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371325; AAL50521.1; -.
DR InterPro; IPR009728; BAAIC N.
DR Pfam; PF06989; BAAIC N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BFBCDB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 3.3e-19;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 10
Q8WXS0
ID Q8WXS0 PRELIMINARY; PRT; 73 AA.
AC Q8WXS0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAAIC 1-4-5-6-8.
GN Name=BAAIC;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
R Mozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50380.1; -.
DR Pfam; PF06989; BAAIC N; 1.
SQ SEQUENCE 73 AA; 7871 MW; 98DBC2E6E6E524A CRC64;

Query Match 37.5%; Score 288; DB 2; Length 73;
Best Local Similarity 91.4%; Pred. No. 1.6e-18;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIEPRYESWTRETESTLTYTDSALPSAATDGPAGGLHAG 58
DB 1 MCGGSRADAIEPRYESWTRETESTLTYTDSALPSAATDGPAGGLHAG 58

RESULT 11
Q8WTP6
ID Q8WTP6 PRELIMINARY; PRT; 54 AA.
AC Q8WTP6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BAAIC isoform 1-8 (BAAIC protein) (BAAIC 1-8).
GN Name=BAAIC;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinenen K.,
R Mozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAAIC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RX STRAUSBERG R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
RA Datchenko L., Marsina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.B.,
RA Brownstein M.J., Uscid T.B., Toshlyuk S., Carninci P., Prange C.,
RA Rana S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultik S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
FAhey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalak U., Smalhus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF371323; AAL50519.1; -;
 DR EMBL; BC035038; AAL50308.1; -;
 DR EMBL; AF363578; AAL50378.1; -;
 DR Pfam; PF06989; BAALC N; 1.
 SQ SEQUENCE 54 AA; 5663 MW; FEF0B4EABED9B825 CRC64;

Query Match 36.2%; Score 278; DB 2; Length 54;
 Best Local Similarity 94.4%; Pred. No. 9e-18;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 54
 DB 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 54

RESULT 12
 ID Q8WXS2 PRELIMINARY; PRT; 80 AA.
 AC Q8WXS2;
 DT 01-JUN-2003 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE BAALC 1-2.
 GN Name=BAALC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heimonen K.,
 RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BAALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF363578; AAL50382.1; -;
 DR Pfam; PF06989; BAALC N; 1.
 SQ SEQUENCE 80 AA; 8573 MW; 70BE9B91C3245E2D CRC64;

Query Match 35.5%; Score 272; DB 2; Length 80;
 Best Local Similarity 94.3%; Pred. No. 5e-17;
 Matches 50; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 53
 DB 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 53

RESULT 13
 ID Q8WNE8 PRELIMINARY; PRT; 54 AA.
 AC Q8WNE8;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE BAALC isoform 1-8.
 GN Name=BAALC;
 OS Sus scrofa (Pig).
 RN [1]

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OK NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heimonen K.,
 RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BAALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF371326; AAL50522.1; -;
 DR Pfam; PF06989; BAALC N; 1.
 SQ SEQUENCE 54 AA; 5681 MW; 6538C3DABED9B825 CRC64;

Query Match 34.8%; Score 267; DB 2; Length 54;
 Best Local Similarity 92.6%; Pred. No. 8.8e-17;
 Matches 50; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 54
 DB 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 54

RESULT 14
 ID Q801V5 PRELIMINARY; PRT; 32 AA.
 AC Q801V5;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE SI:ZC215113.3 (Novel protein similar to human brain and acute
 DE leukemia, cytoplasmic (BAALC)) (Fragment).
 GN Name=SI:ZC215113.3; (Zebrafish)
 OS Brachydanio rerio (Zebrafish)
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OK NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL807244; CAD87801.1; -;
 DR InterPro; IPR009728; BAALC_N.
 DR Pfam; PF06989; BAALC N; 1.
 FT NON_TER 32
 FT NON_TER 32
 SQ SEQUENCE 32 AA; 3663 MW; 916A5445D263B7F7 CRC64;

Query Match 21.0%; Score 161; DB 2; Length 32;
 Best Local Similarity 90.6%; Pred. No. 1.7e-07;
 Matches 29; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 32
 DB 1 MCGGSRADAIERYESWTRETESTWLTYSDDLPSAATDSGPEAGGLHAG 32

RESULT 15
 ID Q6ZMJ6 PRELIMINARY; PRT; 613 AA.
 AC Q6ZMJ6;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE FLJ00420 protein (Fragment).
 GN Name=FLJ00420;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
 OK NCBI_TaxID=9606;
 RN [1]

```

RP SEQUENCE FROM N.A.
RC TISSUE=Spleen;
RA Jikuya H., Takano J., Nomura N., Kikuno R., Nagase T., Ohara O.;
RL Submitted (APR-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK160386; BAD18729.1; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0030528; F:transcription regulator activity; IEA.
DR GO; GO:0045449; P:regulation of transcription; IEA.
DR InterPro; IPR007207; Not N.
DR InterPro; IPR010978; tRNA_binding_arm.
DR Pfam; PF04065; Nct3; 1.
DR NON_TER 1
SQ SEQUENCE 613 AA; 64243 MW; 8D78E87BD58F9E2C CRC64;
OY Query March 12.1%; Score 92.5; DB 2; Length 613;
DB Best Local Similarity 31.9%; Pred. No. 8;
Matches 30; Conservative 7; Mismatches 32; Indels 25; Gaps 3.

31 TDSDALPFAAATDSCPEAGGLH-----AGLTEDGSPNSGVLRPAAP-- 71
||:|||||
216 TDSSEVSOSPAKNGSPVHNSHNPSPAPVPTYPSPGPAPASALSTTPGNNGVPAPAPPS 275
OY 72 --GGIANPEKKNCGTCCPNSSOSLS---SGPLT 99
|:|:|:|:|:|:|:|:|
Db 276 ALGPVASEPAPSHNSTPAPVYQAVAPAPPSGPST 309

```

This Page Blank (uspto)


```

RESULT 1
US-09-949-016-8824
; Sequence 8824, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 8824
; LENGTH: 350
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-8824

Query Match
Best Local Similarity 11.4%; Score 88.5; DB 4; Length 350;
Matches 43; Conservative 15; Mismatches 64; Indels 57; Gaps 10;

QY 2 GCGGSRDAIEPRYESWTRETESTLWTTSDAPPSAAPPDSC-----PEAGL--- 51
Db 21 GLGGSTLD-----ELEWKAKREK---MRAKONPPGAPPGGSSDAAGKPPAGLGRP 71
QY 52 ---HSGMLIEDLPNGVPRSTA---PGSI-----PNEPKTTN-----CETQC 87
Db 72 AAAAANELNNLNP- -GAPAPAPVAVGPGGVCNCAVGSAMLTTRAPGGRRSDEPPAASASNA 130
QY 88 PNPQ-----SLSSGPLTKQONGLOTTTEAKRDAKMPAKEVTINTVDSIQOMD 134
Db 131 PPGQRDEEEDPGVPEKKGSGSPARKKG- -IEKKLREKRSSTGVNIPAAECLDEYE 188

RESULT 2
US-08-773-870-4
; Sequence 4, Application US/08773870
; Patent No. 5912143
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Goli, Surya K.
; TITLE OF INVENTION: NOVEL HUMAN MAGE-LIKE PROTEIN
;

```


US-09-252-991A-25785

Query Match 10.4%; Score 81; DB 4; Length 208;
Best Local Similarity 23.0%; Pred. No. 0.83;
Matches 37; Conservative 18; Mismatches 70; Indels 36; Gaps 4;

```

QY      CGGSR-----ADALEPYYVSMTRRETSTLYLTDDADAPSAAPDGGEAGLHSGMLED 58
Db      CGGSASIRPARARBPSPKASSPTTAMTGTMTMRPSPYSAKAREPTTCGASASPAATP 87
QY      59 GL-----PSNGVPRSTAPGAIPIPEKKTNCETQ-----CPN 89
Db      88 GTTGCTPPPPAPPTKSSAAGGCAMPDCHACRRRTARPPPIPVAPWISTSPASVPGATC 147
QY      90 POSLSGGPLTKONGLOTTAEAKRDAC--RMPAEVITINVT 127
Db      148 PTSSRRGPTSSRRRRTRCTAPTRCSRRYARRAKKATITRCT 188

```

RESULT 6
US-09-94

```

: Sequence 11662, Application US/09949016
: Patent No. 6812339
: GENERAL INFORMATION:
: APPLICANT: VENTER, J. Craig et al.
: TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
: TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
: FILE REFERENCE: CLO01307
: CURRENT APPLICATION NUMBER: US/09/949, 016
: PRIOR FILING DATE: 2000-04-14
: PRIOR APPLICATION NUMBER: 60/241,755
: PRIOR FILING DATE: 2000-10-20
: PRIOR APPLICATION NUMBER: 60/237,768
: PRIOR FILING DATE: 2000-10-03
: PRIOR APPLICATION NUMBER: 60/231,498
: PRIOR FILING DATE: 2000-09-08
: NUMBER OF SEQ ID NOS: 207012
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 11662
: LENGTH: 507
: TYPE: prt
: ORGANISM: Human
: US-09-949-016-11662

```

	Query Match	10.3%	Score 80 ; DB 4 ; Length 507 ;
	Best Local Similarity	29.4% ; Pred. No. 3.7 ;	
Matches	30 ; Conservative	7 ; Mismatches 57 ; Indels 8 ; Gaps 2.	
OY	24 ESTWLTLYDSDAPRPA-----APDSGEAAGLHSGMLIEDGLGNSGVPRSTAGIIPN 76 :: : :		
Dd	181 KSFTLLITVFPNPQVAITYHRAIKTVTDGPREFRRHRKQLTDSKPSLPDRSLDIGRIPIH 240 		
OY	77 PEKTNCETQCENPPOSLSGGPLTKOKNGILQTTEAKRDAAKMP 118 :		
Dd	241 PSMRGVGVPQNRP-R-SLNASAPSPPFGOSQSITTDRKAOSSP 221 :		

```

RESULT 7
US-09-086-663A-81
; Sequence 81, Application US/09086663A
; Patent No. 6518063
; GENERAL INFORMATION:
; APPLICANT: DUCY, PATRICIA
; APPLICANT: KARSENTY, GEARAD
; TITLE OF INVENTION: OSF2/CBPA1 COMPOSITIONS AND METHODS OF USE
; FILE REFERENCE: UTSC:525
; CURRENT APPLICATION NUMBER: US/09/086, 663A
; CURRENT FILING DATE: 1998-05-29
; PRIOR APPLICATION NUMBER: 60/080,189
; PRIOR FILING DATE: 1998-03-24
; PRIOR APPLICATION NUMBER: 60/048,430
; PRIOR FILING DATE: 1997-05-29
; NUMBER OF SEQ ID NOS: 83

```

```

; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 81
; LENGTH: 521
; TYPE: PR1
; ORGANISM: Homo sapiens
US-09-086-663A-81

```

	Query March	10.3%;	Score 80;	DB 4;	Length 521;
	Best Local Similarity	29.4%;	Pred. No. 3.9;		
	Matches	30;	Conservative	7;	Mismatches 57; Indels 8; Gaps 2.
Oy	24 ESTWLTATDSDAPPSSAA-----APDSGPEAGLHSGMLEDGLPNSGVPRSTARGGIPN	76			
Dd	195 KSFLLITTFVFNPPQAVATHRAIKYTDGPREPRRRHKOKLDSDKPSLESDRLSDLGRIPH	254			
Oy	77 PEKTNGCECTCCPNPOSILSSGFPLTKONGKLQTTEAKRDAKMP	118			
Dd	255 PSMRVGVPPQNPR-SLNASAPSPNPQGQSITDPKQAQSPF	295			

RESULT 8
118-09-09

US-03-086-663A-82
 ; Sequence 82, Application US/09086663A
 ; Patent No. 6518063
 ; GENERAL INFORMATION:
 ; APPLICANT: DUCY, PATRICIA
 ; APPLICANT: KARSENTY, GERRARD
 ; TITLE OR INVENTION: OSF2/CBPA1 COMPOSITIONS AND METHODS OF USE
 ; FILE REFERENCE: UTSC:525
 ; CURRENT APPLICATION NUMBER: US/09/086,663A
 ; CURRENT FILING DATE: 1998-05-29
 ; PRIOR APPLICATION NUMBER: 60/080,189
 ; PRIOR FILING DATE: 1998-03-24
 ; PRIOR APPLICATION NUMBER: 60/048,430
 ; PRIOR FILING DATE: 1997-05-29
 ; NUMBER OF SEQ ID NOS: 83
 ; SOFTWARE: PatentIn Ver. 2.1
 ; SEQ ID NO 82
 ; LENGTH: 528
 ; TYPE: PR1
 ; ORGANISM: Mus musculus
 ; US-09-086-663A-82

	Query Match	10.3%	Score 80;	DB 4;	Length 528;
	Best Local Similarity	29.4%;	Pred. No. 3.9;		
	Matches	30;	Conservative	7;	Mismatches 57; Indels 8; Gaps 2.
Oy	24	ESTWLTYYTSDSDAPPSAA-----	APDSGPAGGLGHSMEDGLPSPNVPRSTARGCIPN	76	
	:	: :: :	:: :	:	:
Dd	202	KSFLLITTVFPNPQVATYHRAIKVTYWDGPEPPRRHRKOLDKSPLFSDDLGRILPH		261	
Oy	77	PEKTNCETCCPNFQSLSGCLTKOKNGLQTTEAKRDAKMP		118	
	:	: :: :	:: :	:	:
Dd	262	PSMEVGVPQNFRP-SLNSAPSFPNPGQSQSITTPRQAQSSP		302	

```

, RESULT 9
, US-09-086-663A-71
, Sequence 71, Application US/09086663A
, Patent No. 6518063
, GENERAL INFORMATION:
, APPLICANT: DUCY, PATRICIA
, APPLICANT: KARSENTY, GERRARD
, TITLE OF INVENTION: OSF2/CBPAL COMPOSITIONS AND METHODS OF USE
, FILE REFERENCE: USPC:525
, CURRENT APPLICATION NUMBER: US/09/086, 663A
, CURRENT FILING DATE: 1998-05-29
, PRIOR APPLICATION NUMBER: 60/080,189
, PRIOR FILING DATE: 1998-03-24
, PRIOR APPLICATION NUMBER: 60/048,430
, PRIOR FILING DATE: 1997-05-29
, NUMBER OF SEQ ID NOS: 83
, SOFTWARE: PatentIn Ver. 2.1

```

SEQ ID NO 71
LENGTH: 548
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-086-663A-71

Query Match 10.3%; Score 80; DB 4; Length 548;
Best Local Similarity 29.4%; Pred. No. 4.1; Indels 57; Gaps 2;
Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTSDAPPSAA-----APDSGPEAGLHSGMLDGLPSNGVPRSTAPGGIN 76
DB 202 KSFLLITVFTNPQVATYHRAIKTVDPGRPRRRRQKLDGSKPLFSDDLRIH 261
QY 77 PEKTNCECQCPNPOSLSGPLTOKONGLOTTEAKDKAMP 118
DB 262 PSMRVGVPPQNPRP-SLNSAPSPFNPGQSQITDPRQAQSSP 302

RESULT 10

US-09-086-663A-2
Sequence 2, Application US/09086663A
Patent No. 6518063
GENERAL INFORMATION:
APPLICANT: DUCY, PATRICIA
APPLICANT: KARSENTY, GERARD
TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE
FILE REFERENCE: UTSC:525
CURRENT APPLICATION NUMBER: US/09/086,663A
CURRENT FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/080,189
PRIOR FILING DATE: 1998-03-24
PRIOR APPLICATION NUMBER: 60/048,430
PRIOR FILING DATE: 1997-05-29
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 2
LENGTH: 596
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-086-663A-2

Query Match 10.3%; Score 80; DB 4; Length 596;
Best Local Similarity 29.4%; Pred. No. 4.7;
Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTSDAPPSAA-----APDSGPEAGLHSGMLDGLPSNGVPRSTAPGGIN 76
DB 270 KSFLLITVFTNPQVATYHRAIKTVDPGRPRRRRQKLDGSKPLFSDDLRIH 329
QY 77 PEKTNCECQCPNPOSLSGPLTOKONGLOTTEAKDKAMP 118
DB 330 PSMRVGVPPQNPRP-SLNSAPSPFNPGQSQITDPRQAQSSP 370

RESULT 11

US-09-086-663A-80
Sequence 80, Application US/09086663A
Patent No. 6518063
GENERAL INFORMATION:
APPLICANT: DUCY, PATRICIA
APPLICANT: KARSENTY, GERARD
TITLE OF INVENTION: OSF2/CBFA1 COMPOSITIONS AND METHODS OF USE
FILE REFERENCE: UTSC:525
CURRENT APPLICATION NUMBER: US/09/086,663A
CURRENT FILING DATE: 1998-05-29
PRIOR APPLICATION NUMBER: 60/080,189
PRIOR FILING DATE: 1998-03-24

PRIOR APPLICATION NUMBER: 60/048,430
PRIOR FILING DATE: 1997-05-29
NUMBER OF SEQ ID NOS: 83
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 80
LENGTH: 596
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-09-086-663A-80

Query Match 10.3%; Score 80; DB 4; Length 596;
Best Local Similarity 29.4%; Pred. No. 4.7;
Matches 30; Conservative 7; Mismatches 57; Indels 8; Gaps 2;

QY 24 ESTWLTYYTSDAPPSAA-----APDSGPEAGLHSGMLDGLPSNGVPRSTAPGGIN 76
DB 270 KSFLLITVFTNPQVATYHRAIKTVDPGRPRRRRQKLDGSKPLFSDDLRIH 329
QY 77 PEKTNCECQCPNPOSLSGPLTOKONGLOTTEAKDKAMP 118
DB 330 PSMRVGVPPQNPRP-SLNSAPSPFNPGQSQITDPRQAQSSP 370

RESULT 12

US-09-902-540-14218
Sequence 14218, Application US/09902540
Patent No. 6833447
GENERAL INFORMATION:
APPLICANT: Goldman, Barry S.
APPLICANT: Hinkle, Steven C.
APPLICANT: Wiegand, Roger C.
TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
FILE REFERENCE: 38-10(15849)B
CURRENT APPLICATION NUMBER: US/09/902,540
CURRENT FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: 60/217,883
PRIOR FILING DATE: 2000-07-10
NUMBER OF SEQ ID NOS: 16825
SEQ ID NO 14218
LENGTH: 1036
TYPE: PRT
ORGANISM: Myxococcus xanthus
US-09-902-540-14218

Query Match 10.3%; Score 80; DB 4; Length 1036;
Best Local Similarity 24.5%; Pred. No. 10;
Matches 27; Conservative 14; Mismatches 47; Indels 22; Gaps 3;

QY 31 TDSAPPSAAAPDSGPEAGLHSGMLDGLPSNGVPRSTAPGGINP-----EKTNCET 85
DB 595 SDEASPPGAAVDHGAP-----PSEESPANTAGAVPPPLAPQSECEPATDPT 641
QY 86 QCPNPOSLSGPLTOKONGLOTTEAKDKAMPKAEVINTWDSQOMDR 135
DB 642 APPSPSPAEDAPT---ALTRALDGPVAPDPDGMELQILSDEYSPLE 687

RESULT 13

US-09-948-016-7040
Sequence 7040, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755

;; PRIOR FILING DATE: 2000-10-20
;; PRIOR APPLICATION NUMBER: 60/237,768
;; PRIOR FILING DATE: 2000-10-03
;; PRIOR APPLICATION NUMBER: 60/231,498
;; PRIOR FILING DATE: 2000-09-08
;; NUMBER OF SEQ ID NOS: 207012
;; SOFTWARE: FastSeq for Windows Version 4.0
;; SEQ ID NO 7040
;; LENGTH: 500
;; TYPE: PRT
;; ORGANISM: Human
US-09-949-016-7040

Query Match 10.1%; Score 79; DB 4; Length 500;
Best Local Similarity 26.3%; Pred. No. 4.7;
Matches 30; Conservative 14; Mismatches 50; Indels 20; Gaps 4;

QY 9 DAIEPRYESTWRTRETESTWLTYSDDAPPSAAPDSGPEAGLHSGMLEDGLPSNGVPRS 68
DB 148 DEAESQYDSGISLRL-----RSLPESTASAPASG-----SDGSPQ---PCT 188
QY 69 TABGIRNPEKKTCTCETQCPNPOSISGGLTQKONGLOTTEAKRDAKMPAKY 122
DB 189 HPPGPVXEPQEKEDADGERAD-STYGSSTLYTLTLGGEPEADPPAPRLPLPHV 241

RESULT 14

US-09-252-991A-21479
; Sequence 21479, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 21479
; LENGTH: 330
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21479

Query Match 10.1%; Score 78.5; DB 4; Length 330;
Best Local Similarity 26.2%; Pred. No. 3;
Matches 34; Conservative 13; Mismatches 50; Indels 33; Gaps 6;

QY 2 GCGGSRADAI-----PRYESTWRTRETESTWLTYSDDAPPSA----- 39
DB 26 GCMGSDRTPVQLRAGYAKEVQGRPLGKQAEASATGLSTVPLPSPMRLCGAIVT 85
QY 40 -AAPDSGPEAGLHSGMLEDGLPSNGVPRSTAPGIRNPEKKTN--CETQCPNPOSISG 96
DB 86 DATGGLPRGGGRHRRRRHPSPTHLPRS--PG---KORRLHLRRQAPAP---AAG 137
QY 97 PLTOKONGLO 106
DB 138 PARQDPGLQ 147

RESULT 15

US-09-976-594-427
; Sequence 427, Application US/09976594
; Patent No. 6673549
; GENERAL INFORMATION:
; APPLICANT: Furness, Michael
; APPLICANT: Buchbinder, Jenny
; TITLE OF INVENTION: GENES EXPRESSED IN C3A LIVER CELL CULTURES TREATED WITH STEROIDS

;; FILE REFERENCE: PA-0041 US
;; CURRENT APPLICATION NUMBER: US/09/976,594
;; CURRENT FILING DATE: 2001-10-12
;; PRIOR APPLICATION NUMBER: 60/240,409
;; PRIOR FILING DATE: 2000-10-12
;; NUMBER OF SEQ ID NOS: 1143
;; SOFTWARE: PERL Program
;; SEQ ID NO 427
;; LENGTH: 462
;; TYPE: PRT
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: misc feature
;; OTHER INFORMATION: Incyte ID No. 6673549 2716015CD1
US-09-976-594-427

Query Match 10.1%; Score 78.5; DB 4; Length 462;
Best Local Similarity 28.6%; Pred. No. 4.7;
Matches 30; Conservative 16; Mismatches 32; Indels 27; Gaps 7;

QY 16 YESWTRTESTWLTYSDDAPPSAAPDSGPEAGLHSGMLEDGLPSNGVPRSTAPGIR 75
DB 298 YRMSAEVTS--TYSDEDRPPK--VPPREPLS-----PSNS--RTSPSKSLP 339
QY 76 NPEKKTNCETQ--CPNPOSISGGLTQKONGLOTTEAKRDAKMP 118
DB 340 SYLNGVMPPTQSFAPDPKIVSSKAL--QRQ-----SGSASKVP 377

Search completed: April 12, 2005, 15:43:29
Job time : 27.6667 secs

This Page Blank (uspto)

N:Alternate names: NCAM-180
M:Contains: neural cell adhesion molecule, short domain splice form (NCAM-140)
C:Species: Mus musculus (house mouse)
C:Date: 31-Mar-1993 #sequence revision 31-Mar-1993 #text_change 09-Jul-2004
C:Accession: A29673; S00844; S00384; A28281; A44290; S00383
R:Barthele, D.; Santoni, M.J.; Wille, W.; Ruppert, C.; Chaux, J.C.; Hirsch, M.R.; Fontec
EMBO J. 6, 907-914, 1987
A>Title: Isolation and nucleotide sequence of mouse NCAM cDNA that codes for a Mr 79,000
A:Reference number: A29673; MUID:87246524; PMID:3595563
A:Accession: A29673
A:Molecule type: mRNA
A:Residues: 1-548; 'T', 550-571, 'T', 573-574, 'D', 576-588, 'MPS', 593, 'S', 595-599, 'P', 601, 'I'
A:Cross-references: UNIPROT:P1595; EMBL:X00051; NID:G53342; PIDD:CA68263.1; PID:G53343
R:Santoni, M.J.; Barthele, D.; Barbas, J.A.; Hirsch, M.R.; Steinmetz, M.; Goridis, C.; W
Nucleic Acids Res. 15, 8621-8641, 1987
A>Title: Analysis of cDNA clones that code for the transmembrane forms of the mouse neur
A:Reference number: S00844; MUID:88067687; PMID:3684567
A:Accession: S00844
A:Molecule type: mRNA
A:Residues: 529-809, 1077-1115 <SN>
A:Cross-references: EMBL:X06328; NID:G53322; PIDD:CAA29641.1; PID:G817984
R:Barbas, J.A.; Chaux, J.C.; Steinmetz, M.; Goridis, C.
EMBO J. 7, 625-632, 1988
A>Title: Differential splicing and alternative polyadenylation generates distinct NCAM t
A:Reference number: S00382; MUID:88283628; PMID:3396534
A:Accession: S00384
A:Molecule type: DNA
A:Residues: 642-1115 <BAR>
A:Cross-references: EMBL:X07195
R:Barthele, D.; Vopper, G.; Wille, W.
Nucleic Acids Res. 16, 4217-4225, 1988
A>Title: NCAM-180, the large isoform of the neural cell adhesion molecule of the mouse,
A:Reference number: A28281; MUID:88247737; PMID:2454455
A:Accession: A28281
A:Molecule type: mRNA
A:Residues: 804-1081 <BA3>
A:Cross-references: EMBL:X07244; NID:G53321; PIDD:CAA30230.1; PID:G929720
R:Rougon, G.; Marshak, D.R.
J. Biol. Chem. 261, 3396-3401, 1986
A>Title: Structural and immunological characterization of the amino-terminal domain of n
A:Reference number: A44290; MUID:86140120; PMID:3512556
A:Accession: A44290
A:Molecule type: protein
A:Residues: 20-36 <ROU>
C:Comment: NCAM mediates cell-cell adhesion via homophilic binding with another NCAM mol
C:Comment: Several forms of NCAM are produced by alternative splicing. See also PIR:IMS
C:Genetics:
A:Gene: NCAM
A:Map position: 9
A:Introns: 643/3; 701/1; 770/2; 809/2; 1076/2
C:Superfamily: neural cell adhesion molecule; fibronectin type III repeat homology; immu
C:Keywords: alternative splicing; brain; cell adhesion; duplication; heparin binding; si
F:1-19/Domain: signal sequence #status predicted <SIG>
F:20-1115/Product: neural cell adhesion molecule, long domain splice form #status experi
F:20-809,1077-1115/Product: neural cell adhesion molecule, short domain splice form #sta
F:20-711/Domain: extracellular #status predicted <EXT>
F:134-98/Domain: immunoglobulin homology <IMM>
F:132-191/Domain: immunoglobulin homology <IMM2>
F:152-156/Region: heparin binding #status predicted
F:161-165/Region: heparin binding #status predicted
F:228-290/Domain: immunoglobulin homology <IMM3>
F:262-272/Region: NCAM binding #status predicted
F:323-388/Domain: immunoglobulin homology <IMM4>
F:420-482/Domain: immunoglobulin homology <IMM5>
F:519-596/Domain: fibronectin type III repeat homology <FN3>
F:625-688/Domain: fibronectin type III repeat homology <FN3B>
F:712-729/Domain: transmembrane #status predicted <TM>
F:730-1115/Domain: intracellular #status predicted <INT>
F:41-96,139-189,235-288,330-388,427-480/Diulfide bonds: #status predicted
F:222,316,348,424,450,479/Binding site: carbohydrate (asn) #status predicted

```

Matches      35:  Conservative      15:  Mismatches      59:  Indels      23:  Gaps      6:

QY      6  SRADAIIPRYESWTRETESTLWLTYYTDSADAPSPSAAPDSCPEPAGGLHSGMLDEGLPSNGV 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      830  TNSDITLITETFAIAQNSPTSET-TTLTSSIAIPPAVTVPDSSNSVPAG-----CATPSKV 881
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      66  -PRSTAGGIPNPE---KTNCCETQCPNPQSLSS-----GPLTKQNGIQTTEAK 111
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      882  TASSSSPSAPAKVAPLVLDSDTPTSPASPSANSLSTVLANQAVLSPSTPASAG-ETSKAP 940
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      112  RDAKRPAPKEVT 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      941  PAKTASPAPIPT 952

RESULT 11
T35985
probable large Pro/Ala/Gly-rich protein - Streptomyces coelicolor
C:Species: Streptomyces coelicolor
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C:Accession: T35985
R:Murphy, L.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.
submitted to the EMBL Data Library, February 1999
A:Reference number: Z21589
A:Accession: T35985
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-1366 <MUR>
A:Cross-references: UNIPROT:Q9Z529; EMBL:AL035559; PDB: CAB3472.1; GSPDB: GN00070; SCOPED
A:Experimental source: strain A3(2)
C:Genetics:
A:Gene: SCOEDB:SC9F2.06c

Query Match      10.6%; Score 82.5; DB 2; Length 1366;
Best Local Similarity 30.8%; Pred. No. 51;
Matches 33; Conservativity 12; Mismatches 25; Indels 37; Gaps 8;

QY      2  GCGGSRDAIIPRYESWTRETESTLWLTYYTDSADAPSAAPD--SGP-----EAGGLHG- 53
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      458  GAGGAAR-----PTWAKEAPS-----PPASTAPBPWMSGPGGAVGAGGVSSP 499
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

QY      54  --GMLBDG-LPSNGVPR-----STAPGGINPEKKTNCTNCQCPMP 90
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
DB      500  GTGLRGSGTVAAGGVPLFGARGVSGAAGGVPLPTVFTASEPR-PGP 545
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 12
S22571
integrase-like protein FE65 - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 31-Dec-1993 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C:Accession: S22571; S22572
R:Dullio, A.; Zambirano, N.; Mogavero, A.R.; Ammendola, R.; Cimino, F.; Russo, T.
Nucleic Acids Res. 19, 5269-5274, 1991
A:Title: A rat brain mRNA encoding a transcriptional activator homologous to the DNA bin
A:Reference number: S22571; MUID:92020215; PMID:1923810
A:Accession: S22571
A:Molecule type: mRNA
A:Residues: 1-499 <DU11>
A:Cross-references: UNIPROT:Q99MK3; EMBL:X60463; NID:g57559; PDB: CAA42999.1; PID:g57560
A:Accession: S22572
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 31-318 <DU12>
A:Cross-references: EMBL:X60463; NID:g57561; PDB: CAA42998.1; PID:g117617
A>Note: This sequence was submitted to the EMBL Data Library, July 1991
C:Genetics:
A:introns: 88/3; 107/3; 135/3; 157/3; 207/3; 250/2; 252/2; 290/3
C:Keywords: transcription regulation
F:42-78/Domain: WW repeat homology <WW1>

Query Match      10.5%; Score 81.5; DB 2; Length 499;
Best Local Similarity 25.3%; Pred. No. 19;

```

Matches 40; Conservative 20; Mismatches 67; Indels 31; Gaps 9;

QY 13 PRYSES-----WTR--ETESTWL-----TYTSDAPPSAAPPSC--PE-----AGG 50
 Db 35 PNAFETDSDLPAGMMRQODTSGTYWHI PTGTTQWEPGRASPQGNSPQESQLTWTF 94
 QY 51 LHSGLMEDGLPSNGVPRSTAP--GIPNPEKTNCFQCP--NPQSSSGPLTQKONGLOT 107
 Db 95 AHGEGPEGEFWMDEBEEAPMELGLKDPREGT-----LFPSSQSSPEFVPEBENTLPQ 149
 QY 108 TEAKRDAKMPAKEV-TINVTDSIQMDRSRRITKNCV 144
 Db 150 RNANPGIKCFVRSICGVMETEEELAPGRSSVAVNNCI 187

RESULT 13

B90784
 hypothetical protein Eccl242 [imported] - Escherichia coli (strain O157:H7, substrain EDL5
 C/Species: Escherichia coli
 C/Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
 C/Accession: B90784
 R/Hayashi, T.; Makino, K.; Ohnishi, M.; Kurckawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
 gaeawata, N.; Yaenunga, T.; Kuhnara, S.; Shida, T.; Hattori, M.; Shinagawa, H.
 DNA Res. 8, 11-22, 2001
 A/Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and gene
 A/Reference number: A99629; MUID:21156231; PMID:11258796
 A/Accession: B90784
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-2793 <HAY>
 A/Cross-references: UNIPROT:Q8X2Q2; GB:BA000007; PIDN:BAR34665.1; PID:G13360702; GSPDB:C
 C/Genetics: A/Experimental source: strain O157:H7, substrain R1MD:0509952
 A/Genetic: Eccl242

Query Match 10.5%; Score 81.5; DB 2; Length 2793;
 Best Local Similarity 27.3%; Pred. No. 1.4e+02;
 Matches 30; Conservative 15; Mismatches 54; Indels 11; Gaps 4;

QY 35 APPSAAAPDSGPPAGGLHSGMLDGLPSNGVPRSTAP--GIPNPEKTNCFQCPNPQS 92
 Db 791 APENNA-----GROGETLEGDMVRGLPSPDAONATAVREGLPAPDIARN--VRMPQPS 844
 QY 93 LSSGPLTQKONGLOTTEAKRDA---KMPAKEVTINVTDSIQMDRSRR 139
 Db 845 LPRTVRDSLPBLAQAEVRROAGNRDIPQETIAPSESTVSTDEATV 894

RESULT 14

D85644
 hypothetical protein Z1495 [imported] - Escherichia coli (strain O157:H7, substrain EDL5
 C/Species: Escherichia coli
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C/Accession: D85644
 R/Ferna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
 Miller, L.; Grotbeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
 Nature 409, 529-533, 2001
 A/Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
 A/Reference number: A85480; MUID:21074935; PMID:11206551
 A/Accession: D85644
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-2806 <STO>
 A/Cross-references: UNIPROT:Q8X470; GB:AE005174; NID:G12514354; PIDN:AAG55616.1; GSPDB:C
 A/Experimental source: strain O157:H7, substrain EDL933
 C/Genetics: A/Genetic: Z1495

Query Match 10.5%; Score 81.5; DB 2; Length 2806;
 Best Local Similarity 27.3%; Pred. No. 1.4e+02;
 Matches 30; Conservative 15; Mismatches 54; Indels 11; Gaps 4;

QY 35 APPSAAAPDSGPPAGGLHSGMLDGLPSNGVPRSTAP--GIPNPEKTNCFQCPNPQS 92

Db 804 APENNA-----GROGETLEGDMVRGLPSPDAONATAVREGLPAPDIARN--VRMPQPS 857

QY 93 LSSGPLTQKONGLOTTEAKRDA---KMPAKEVTINVTDSIQMDRSRR 139
 Db 858 LPRTVRDSLPBLAQAEVRROAGNRDIPQETIAPSESTVSTDEATV 907

RESULT 15

A85359
 translation initiation factor-1 like protein [imported] - Arabidopsis thaliana
 C/Species: Arabidopsis thaliana (mouse-ear cress)
 C/Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
 C/Accession: A85359
 R/Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999
 A/Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.
 A/Reference number: A85001; MUID:20083488; PMID:10617198
 A/Accession: A85359
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-263 <STO>
 A/Cross-references: UNIPROT:Q9SUH8; GB:NC_001268; NID:G7269969; PIDN:CAB79786.1; GSPDB:C
 C/Genetics: A/Genetic: A85359
 A/Map position: 4

Query Match 10.3%; Score 80.5; DB 2; Length 263;
 Best Local Similarity 28.1%; Pred. No. 11;
 Matches 38; Conservative 13; Mismatches 57; Indels 27; Gaps 7;

QY 2 GCGSGADALIEPRVYSWTRFESTLTYTDSAPPSAAPPDSGPPAGGLHSGMLDGL 59
 Db 15 GRGSGRRFA--PRF-----TLSSSDLTJNGDASFVKGSG-----GLNDRPS 57
 QY 60 --LPSNGVPRSTAPGIGIPNPEKTNCFQCPNPQSSSGPLTQKONGLOTTEAKRDAKRM 117
 Db 58 ALVQNG---SQCPKVPSPPTROT-VKPKPQGPQVAVAPT--TSLNLTVELSRKTNLS 111
 QY 118 PAKETINVTDSIQ 132
 Db 112 LEEYFVNRLLDEALQ 126

Search completed: April 12, 2005, 15:44:45
 Job time : 23.6667 secs

this Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:16:23 ; Search time 94.6667 Seconds

(without alignments)
784.347 Million cell updates/sec

Title: US-10-705-716A-4

Perfect score: 778

Sequence: 1 MGCGSRAADAIERYRESWT.....VTDSIQMDRSRRITKNCVN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: uniprot_sprot:*
2: uniprot_trcemb1:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	778	100.0	145	2 Q9HA93	Q9HA93 homo sapien
2	750.5	96.5	180	2 Q8WA93	Q8WA93 homo sapien
3	664	85.3	145	2 Q8WNE9	Q8WNE9 sus scrofa
4	649	83.4	145	2 Q8VH1	Q8VH1 mus musculus
5	645	82.9	145	2 Q920K5	Q920K5 rattus norv
6	569.5	73.2	149	2 Q8WXS1	Q8WXS1 homo sapien
7	381	49.0	123	2 Q9CVS9	Q9CVS9 mus musculus
8	309	39.7	73	2 Q8WXS0	Q8WXS0 homo sapien
9	299	38.4	54	2 Q8WTP6	Q8WTP6 homo sapien
10	293	37.7	80	2 Q8WXS2	Q8WXS2 homo sapien
11	282	36.2	54	2 Q8WNE8	Q8WNE8 sus scrofa
12	278	35.7	54	2 Q8VBS8	Q8VBS8 mus musculus
13	278	35.7	54	2 Q920N3	Q920N3 rattus norv
14	161	20.7	32	2 Q801V5	Q801V5 brachydanio
15	97.5	12.5	306	2 Q72H02	Q72H02 thermus tho
16	95.5	12.3	1097	1 S24C ARATH	Q9M291 arabidopsis
17	94.5	12.1	1625	2 O55597	O55597 garlic viru
18	93	12.0	291	2 Q6C2R4	Q6C2R4 yarrowia li
19	92.5	11.9	717	2 Q6FKP2	Q6FKP2 candida gla
20	92.5	11.9	1001	1 P5PA RAT	Q9JME1 rattus norv
21	92	11.8	221	2 Q8GAA3	Q8GAA3 archibacte
22	91.5	11.8	222	2 Q8LE41	Q8LE41 arabidopsis
23	91.5	11.8	249	2 Q92X85	Q92X85 bacterioph
24	91.5	11.8	2042	2 Q767L8	Q767L8 sus scrofa
25	91	11.7	189	2 Q73W33	Q73W33 mycobacteri
26	90.5	11.6	670	2 Q828T2	Q828T2 streptomyc
27	90.5	11.6	330	2 Q6BXP9	Q6BXP9 debrayomyce
28	90	11.6	219	2 Q82A55	Q82A55 streptomyc
29	89	11.4	321	2 Q9VUD8	Q9VUD8 drosophila
30	89	11.4	629	2 Q87V71	Q87V71 pseudomonas
31	89	11.4	805	2 Q758W4	Q758W4 ashbya goss

RESULT 1	Q9HA93	PRELIMINARY;	PRT;	145 AA.
AC	Q9HA93			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DT	25-OCT-2004 (TREMBLrel. 28, Last annotation update)			
DE	Hypothetical protein FLJ12015 (BAAUC isoform 1-6-8) (Brain and acute leukemia, cytoplasmic) (BAAUC 1-6-8).			
GN	Name=BAAUC;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
OX	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Whole embryo;			
RX	PubMed=14702039; DOI=10.1038/ng1285;			
RA	Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,			
RA	Wakamatsu A., Hayashi K., Sato H., Negai K., Kimura K., Makita H.,			
RA	Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,			
RA	Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahara K.,			
RA	Marukami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,			
RA	Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,			
RA	Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,			
RA	Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,			
RA	Ninomiyama K., Iehibashi T., Yamashita H., Murakawa K., Fujimori K.,			
RA	Tanai H., Kimata M., Watanabe M., Hirao K., Chiba Y., Iehida S.,			
RA	Ono Y., Takiguchi S., Watanabe S., Yoisida M., Hoshida T., Kusano J.,			
RA	Kanehori K., Takahashi-Fujii A., Hara H., Tanase T., Nomura Y.,			
RA	Togiyama S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,			
RA	Musaehino K., Yuuki H., Oshima A., Sasaki N., Aotsuka S.,			
RA	Yoshikawa Y., Matsumawa H., Ichihara T., Shihata N., Sano S.,			
RA	Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,			
RA	Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,			
RA	Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,			
RA	Yamagaki M., Watanabe K., Kamagai A., Itakura S., Fukunumi Y.,			
RA	Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,			
RA	Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Okumori Y.,			
RA	Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikemori Y., Okamoto S.,			
RA	Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,			
RA	Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,			
RA	Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,			
RA	Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,			
RA	Okumura K., Nagase T., Nomura N., Kikkawa H., Masuho Y., Yamashita R.,			
RA	Nakai K., Yada T., Nakamura Y., Ohara O., Isozaki T., Sugano S.,			
RT	"Complete sequencing and characterization of 21,243 full-length human			
RT	CDNAs."			
RL	Nat. Genet. 36:40-45(2004).			
RN	[2]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;			
RA	Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Helonen K.,			
RA	Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,			

32	89	11.4	1064	2 Q8DIP1	Q8DIP1 synechococ
33	88.5	11.4	340	2 Q96IZ0	Q96IZ0 homo sapien
34	88.5	11.4	690	2 Q9FX17	Q9FX17 arabidopsis
35	88	11.3	302	2 Q93Z47	Q93Z47 arabidopsis
36	88	11.3	472	2 Q8UVG3	Q8UVG3 gallus gall
37	88	11.3	472	2 Q8UWC8	Q8UWC8 gallus gall
38	88	11.3	1194	2 Q6C4Z0	Q6C4Z0 yarrowia li
39	87.5	11.2	192	2 Q67ZP7	Q67ZP7 arabidopsis
40	87.5	11.2	222	2 Q93V72	Q93V72 arabidopsis
41	87.5	11.2	340	2 Q6FHY9	Q6FHY9 homo sapien
42	87.5	11.2	369	2 MAGA HUMAN	P43363 homo sapien
43	87.5	11.2	754	2 Q7TIF0	Q7TIF0 brachydanio
44	87.5	11.2	758	2 Q7SDA2	Q7SDA2 neurospora
45	87.5	11.2	860	2 Q9C982	Q9C982 arabidopsis

ALIGNMENTS

RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
R Klausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,
R Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
R Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
R Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
R Stapleton M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
R Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
R Raha S.S., Loughellano N.A., Peters G.J., Abramson R.D., Mullaney S.J.,
R Bosak S.A., McMan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
R Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
R Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
R Fahey J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,
R Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
R Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
R Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S.,
R Krzywinski M.I., Skalska U., Smalusz D.E., Schnerch A., Schein J.E.,
R Jones S.J., Matra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK022077; BAB13960.1; -
DR EMBL; AF371319; AAL50515.1; -
DR EMBL; BC011517; AAL1517.1; -
DR EMBL; AF363578; AAL50377.1; -
DR InterPro; IPR009728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 1551 MW; CFB92BBE83DD92E CRC64;

Query Match 100.0%; Score 778; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 3, 7e-55;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRTETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDGL 60
DB 1 MCGGSRADAIIPRYESWTRTETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGGINPPEKKTNCETOCNPPOSLSGPTLQKNGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGGINPPEKKTNCETOCNPPOSLSGPTLQKNGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 2
Q8WXS3 PRELIMINARY; PRT; 180 AA.
AC Q8WXS3;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE BAALC 1-5-6-8.
OS Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.

RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heinonen K.,
RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50379.1; -
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 180 AA; 19224 MW; 380183E0F188F684 CRC64;

Query Match 96.5%; Score 750.5; DB 2; Length 180;
Best Local Similarity 80.6%; Pred. No. 7, 8e-53;
Matches 145; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MCGGSRADAIIPRYESWTRTETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDGL 53
DB 1 MCGGSRADAIIPRYESWTRTETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDGL 53
QY 54 -----GMLDGLPSNGVPRSTAPGGINPPEKKTNCET 85
DB 61 KTKAFTDSVSDGLSASAMAPLAFPSHGMLEGLPSNGVPRSTAPGGINPPEKKTNCET 120
QY 86 QCPNPQSSGPTLQKNGLOTTEAKRDARMPAKVITNTDSIQMDRSRRITKNCVN 145
DB 121 QCPNPQSSGPTLQKNGLOTTEAKRDARMPAKVITNTDSIQMDRSRRITKNCVN 180

RESULT 3
Q8WNE9 PRELIMINARY; PRT; 145 AA.
AC Q8WNE9;
DT 01-MAR-2002 (TREMblrel. 20, Created)
DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
DE BAALC isoform 1-6-8.
GN Name=BAALC;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxId=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heinonen K.,
RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371322; AAL50518.1; -
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 15401 MW; C75ED7D00BF82E26 CRC64;

Query Match 85.3%; Score 664; DB 2; Length 145;
Best Local Similarity 84.1%; Pred. No. 5, 6e-46;
Matches 122; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIPRYESWTRTETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDGL 60
DB 1 MCGGSRADAIIPRYESWTRTETSTWLTYSDDAPPSAAPPDGPAGLHSGMLEDGL 60
QY 61 PSNGVPRSTAPGGINPPEKKTNCETOCNPPOSLSGPTLQKNGLOTTEAKRDARMPAK 120
DB 61 PSNGVPRSTAPGGINPPEKKTNCETOCNPPOSLSGPTLQKNGLOTTEAKRDARMPAK 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 4
Q8VHV1

ID Q8VHVL PRELIMINARY; PRT; 145 AA.
AC Q8VHVL;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DE 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE BAAIC isoform 1-6-8 (Mus musculus 16 days neonate cerebellum cDNA,
DE RIKEN full-length enriched library, clone:965002H16 product:brain and
DE acute leukemic, cytoplasmic, full insert sequence).
GN Name=BaalC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12986/SEVITAC;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heimonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BaalC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=99279233; PubMed=10349686; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX The PANTOM Consortium;
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Komori H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Komori H., Akiyama J., Nishi K., Katsunari T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwaga K.,
RA Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohata E., Watanabe K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsunaga S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,

RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katoh H., Kawai J., Kojima Y., Kondo S., Komori H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toyota T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (Apr-2002) to the EMBL/Genbank/DBD databases.
DR EMBL; AF371320; AAL50516.1; -
DR EMBL; AK079337; BAC37611.1; -
DR MGI; MGI:1928704; BaalC.
DR pfam; PF06989; BaalC_N; 1.
SQ SEQUENCE 145 AA; 15515 MW; 4972670A618C4D6D CRC64;
Query Match 83.4%; Score 649; DB 2; Length 145;
Best Local Similarity 83.4%; Pred. No. 9, 1e-45;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
QY 1 MCCGSRADAIIPRYESWTRETESTWLTYSDAAPPSAAADSGPEAGLHSGMLDGL 60
DB 1 MCCGSRADAIIPRYESWTRETESTWLTYSDAAPPSAAADSGPEAGLHSGMLDGL 60
QY 61 PANGVPRSTAPGCIENPEKTKNCETQCPNPQSLSSGFLTKQKNGLTTEAKRDAKMPAK 120
DB 61 SNGVLRPAAPGAIANPEKKNCGTQCPNSQNLSSGFLTKQKNGLTTEAKRDAKMSAR 120
QY 121 EVTINVTDSIQMDRSRRITKNCVN 145
DB 121 EVTINVTENIRQMDRSKRTKNCIN 145
RESULT 5
Q920K5 PRELIMINARY; PRT; 145 AA.
ID Q920K5;
AC Q920K5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 01-DEC-2001 (TrEMBLrel. 19, Last sequence update)
DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Dem-A20-4 (BAAIC isoform 1-6-8).
GN Name=Dem-A20-4; Synonyms=BaalC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang X., Tian Q., Li W., Okano A., Suzuki T.;
RL Submitted (Oct-2001) to the EMBL/Genbank/DBD databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heimonen K.,
RA Mrozek K., Sill H., Knuttila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BaalC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AB073318; BAB70507.1; -
DR EMBL; AF371321; AAL50517.1; -
DR Pfam; PF06989; BaalC_N; 1.
SQ SEQUENCE 145 AA; 15475 MW; D5A27AD67456F341 CRC64;
Query Match 82.9%; Score 645; DB 2; Length 145;
Best Local Similarity 83.4%; Pred. No. 1, 9e-44;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
QY 1 MCCGSRADAIIPRYESWTRETESTWLTYSDAAPPSAAADSGPEAGLHSGMLDGL 60
DB 1 MCCGSRADAIIPRYESWTRETESTWLTYSDAAPPSAAADSGPEAGLHAGVLEDP 60

```

QY 61 PSNGVPRSTAGGIPNPEKKTNCETQCPNQSISGGLTQKONGLOTTEAKRDAKMPAK 120
DB 61 SSGVLRPAAPGSIAPPEKKNCGTQCPNQSISGGLTQKONGLOTTEAKRDAKMSAR 120
QY 121 EYINTVTDSTQCDRSRRITKNCVN 145
DB 121 EVAISVTENIRQMDRSKRVTKNCIN 145

RESULT 6
Q8WXS1 PRELIMINARY; PRT; 149 AA.
AC Q8WXS1;
DB 01-MAR-2002 (T-EMBLrel. 20, Created)
DB 01-MAR-2002 (T-EMBLrel. 20, Last sequence update)
DB 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DB BAALC.1-5-6-7-8.
GN Name=BAALC;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1];
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinenen K.,
RA Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50381.1; -.
DR Gene; GNCN:14333; BAALC.
PRFam: P06989; BAALC.N:1.
SQ SEQUENCE 149 AA; 15577 MW; DB26C19969B91D6E CRC64;

Query Match 73.2%; Score 569.5; DB 2; Length 149;
Best Local Similarity 75.7%; Pred. No. 2.4e-38;
Matches 109; Conservative 0; Mismatches 0; Indels 35; Gaps 1;

QY 1 MCGGSGRAADIERRYESWTRETSWTLYTTDSDAPPSAAPPSCGEGAGLHS----- 53
DB 1 MCGGSGRADAIERYESWTRETSWTLYTTDSDAPPSAAPPSCGEGAGLHSVLEAKS 60
QY 54 -----GMLDGLPSNGVPRSTAGGIPNPEKKTNCET 85
DB 61 KIKAPDVSDEGLFSAASKAPLAIVSHKLEBDGLPSNGVPRSTAGGIPNPEKKTNCET 120
QY 86 QCPNQSISGGLTQKONGLOTTE 109
DB 121 QCPNQSISGGLTQKONGLOTTE 144

RESULT 7
Q9CY99 PRELIMINARY; PRT; 123 AA.
AC Q9CY99;
DB 01-JUN-2001 (T-EMBLrel. 17, Created)
DB 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DB 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DB Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
DB enriched library, clone:2810457D07 product:brain and acute leukemia,
DB cytoplasmic, full insert sequence.
GN Name=BaalC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1];
RP SEQUENCE FROM N.A.
RX STRAIN=C57BL/6J; TISSUE=whole body;
RC MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
```

```

RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RA The FANTOM Consortium;
RT the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
RN [4];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Alzawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Kono H., Akiyama J., Nishi K., Kitsuana T., Tashiro H., Itoh M.,
RA Sumi N., Iehi Y., Nakamura S., Hazama M., Nishine T., Harada T.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kasahagi K.,
RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Watanuki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6];
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=whole body;
RA Adachi T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA Hangaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA Imocani K., Iehi Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA Kawai J., Kojima Y., Kono H., Kouda M., Koya S., Kurihara C.,
RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA Muramatsu M., Hayashizaki Y.;
RT Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK013358; BAB28808.1; -.
DR MGD; MGI:1928704; BaalC.
SQ SEQUENCE 123 AA; 13391 MW; 33DFC85F8CFA0AB CRC64;

Query Match 49.0%; Score 381; DB 2; Length 123;
Best Local Similarity 77.4%; Pred. No. 2.9e-23;
Matches 72; Conservative 9; Mismatches 12; Indels 0; Gaps 0;

QY 53 SGMLEDGIPSNVPRSTAGGIPNPEKKTNCETQCPNQSISGGLTQKONGLOTTEAKR 112
DB 31 SGVLEDGSSNGVLRPAAPGSIAPPEKKNCGTQCPNQSISGGLTQKONGLOTTEAKR 90
QY 113 DAKRMPAKVINTVTDSTQCDRSRRITKNCVN 145
DB 91 DAKRMSAREVAINTENIRQMDRSKRVTKNCIN 123

RESULT 8
```



```

Q8MXS0
ID O8MXS0 PRELIMINARY; PRT; 73 AA.
AC O8MXS0;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BAALC 1-4-5-6-8.
Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50380.1; -.
SQ SEQUENCE 73 AA; 7871 MW; 98DBC2E6E6F524A CRC64;

Query Match 39.7%; Score 309; DB 2; Length 73;
Best Local Similarity 96.6%; Pred. No. 9.8e-18;
Matches 56; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSGMED 58
Db 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSGCLEE 58

RESULT 9
Q8WTP6 PRELIMINARY; PRT; 54 AA.
AC O8WTP6;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE BAALC isoform 1-8 (BAALC protein) (BAALC 1-8).
Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50380.1; -.
SQ SEQUENCE 54 AA; 5663 MW; FEF0B4EABED9B829 CRC64;

Query Match 38.4%; Score 299; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.4e-17;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 54

RESULT 10
Q8MXS2 PRELIMINARY; PRT; 80 AA.
AC O8MXS2;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BAALC 1-2.
Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50382.1; -.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 80 AA; 8573 MW; 70BE9B91C3245E2D CRC64;

Query Match 37.7%; Score 293; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 53
Db 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 53

RESULT 11
Q8WNE8 PRELIMINARY; PRT; 54 AA.
AC O8WNE8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BAALC isoform 1-8.
Name=BAALC;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

```

RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywicki M.I., Skalska U., Smailus D.E., Schmeck A., Schein J.E.,
RA Jones S.J., Mitra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA Strausberg R.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF371323; AAL50519.1; -.
DR EMBL; BC035038; AAL35038.1; -.
DR EMBL; AF363578; AAL50378.1; -.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5663 MW; FEF0B4EABED9B829 CRC64;

Query Match 38.4%; Score 299; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 4.4e-17;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 54

RESULT 10
Q8MXS2 PRELIMINARY; PRT; 80 AA.
AC O8MXS2;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BAALC 1-2.
Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
RA Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
RT lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50382.1; -.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 80 AA; 8573 MW; 70BE9B91C3245E2D CRC64;

Query Match 37.7%; Score 293; DB 2; Length 80;
Best Local Similarity 100.0%; Pred. No. 2.1e-16;
Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 53
Db 1 MCGGSRADAIERYYSWTRETSTWLTYYTDSAPPSAAPPDGPAGGLHSG 53

RESULT 11
Q8WNE8 PRELIMINARY; PRT; 54 AA.
AC O8WNE8;
DT 01-MAR-2002 (TREMBlrel. 20, Created)
DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE BAALC isoform 1-8.
Name=BAALC;
OS Sus scrofa (Pig).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
OX NCBI_TaxID=9823;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinenen K.,
R Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371326; AAL50522.1; -.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; 6538C3DABED9B825 CRC64;

Query Match 36.2%; Score 282; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 1e-15;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 54
|||||

RESULT 12
Q8VBS8 PRELIMINARY; PRT; 54 AA.
ID 08VBS8
AC 08VBS8;
DT 01-MAR-2002 (TRENBLrel. 20, Created)
DT 01-MAR-2002 (TRENBLrel. 20, Last sequence update)
DE 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE BAALC isoform 1-8.
GN Name=BaalC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=12956/Syvatag;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinenen K.,
R Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371324; AAL50520.1; -.
DR Pfam; PF06989; BAALC_N; 1.
DR MGD; MGI:1928704; BaalC.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0EBFBCDB829 CRC64;

Query Match 35.7%; Score 278; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 2.2e-15;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 54
|||||

RESULT 13
Q790N3 PRELIMINARY; PRT; 54 AA.
ID 0790N3
AC 0790N3;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE BAALC isoform 1-8.
GN Name=BaalC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;

OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=SDRague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinenen K.,
R Mrozek K., Sill H., Knuttila S., Kolitz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371325; AAL50521.1; -.
DR InterPro; IPR009728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0EBFBCDB829 CRC64;

Query Match 35.7%; Score 278; DB 2; Length 54;
Best Local Similarity 94.4%; Pred. No. 2.2e-15;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 54
Db 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 54
|||||

RESULT 14
Q801V5 PRELIMINARY; PRT; 32 AA.
ID 0801V5
AC 0801V5;
DT 01-JUN-2003 (TRENBLrel. 24, Created)
DT 01-JUN-2003 (TRENBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE SI:ZC215113.3 (Novel protein similar to human brain and acute
leukemia, cytoplasmic (BAALC)) (Fragment).
GN Name=SI:ZC215113.3;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OX NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RA Sehba H.;
RL Submitted (JUN-2003) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL807244; CAD87801.1; -.
DR InterPro; IPR009728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
FT NON_TER 32
SQ SEQUENCE 32 AA; 3663 MW; 916A5445D263B7F7 CRC64;

Query Match 20.7%; Score 161; DB 2; Length 32;
Best Local Similarity 90.6%; Pred. No. 3.1e-06;
Matches 29; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 32
Db 1 MCGGSRADAIERYESWTRETESTWLTYYTDSADPSPAAPDGPAGGLHSG 32
|||||

RESULT 15
Q72H02 PRELIMINARY; PRT; 306 AA.
ID 072H02
AC 072H02;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DT 05-JUL-2004 (TRENBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TRENBLrel. 27, Last annotation update)
DE Hypothetical protein.
GN OrderedNames=ITC1434;
OS Thermus thermophilus (strain HB27 / ATCC BAA-163 / DSM 7039).
OC Bacteria; Deinococcus-Thermus; Deinococci; Thermales; Thermaceae;
OC Thermus.
OX NCBI_TaxID=262724;
RN [1]

RP SEQUENCE FROM N.A.
 RX PubMed-15064768;
 RA Henne A., Brueggemann H., Raasch C., Wlezer A., Hartach T.,
 RA Liebegang H., Johann A., Lienard T., Gohl O., Martinez-Arias R.,
 RA Jacobi C., Starkuviene V., Schlenczek S., Dencker S., Huber R.,
 RA Klenk H.-P., Kramer W., Merkl R., Gottschalk G., Fritz H.-J.,
 RT "The genome sequence of the extreme thermophile Thermus
 thermophilus."
 RL Nat. Biotechnol. 22:547-553(2004).
 DR EMBL; AB017305; AAS81776.1; -.
 KW Complete proteome.
 SQ SEQUENCE 306 AA; 31927 MM; 6184FB203F4CE1FA CRC64;

Query Match 12.5%; Score 97.5; DB 2; Length 306;
 Best Local Similarity 30.6%; Pred. No. 5.6;
 Matches 26; Conservative 12; Mismatches 38; Indels 9; Gaps 2;
 QY 36 PPSAAPPDSCPEAGLHSGMLEDLPSNGVPRSTAPGIPNPEKTNCE---TQCPNPQ 91
 DB 164 PPSAAPPDSCPEAGLHSGMLEDLPSNGVPRSTAPGIPNPEKTNCE---TQCPNPQ 91
 QY 92 SLSSGP-----LTQKONGLOTTEAK 111
 DB 224 ASRRGPGAGGTLREKEGALPTQAO 248

Search completed: April 12, 2005, 15:42:10
 Job time : 96.6667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:15:13 ; Search time 103.667 Seconds
(without alignments)
540.967 Million cell updates/sec

Title: US-10-705-716a-8

Perfect score: 767
Sequence: 1 MCGGSRADAIERYRYESWT.....VTENINQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: A_Geneseq_16Dec04:*

1: geneseqp19808:*\n2: geneseqp19908:*\n3: geneseqp20008:*\n4: geneseqp20018:*\n5: geneseqp20028:*\n6: geneseqp20038:*\n7: geneseqp20048:*\n8: geneseqp20058:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	767	100.0	145	8	AD048479 Mouse PTH
2	746	97.3	145	8	AD048473 Rat PTH r
3	649	84.6	145	4	AAB95018 Human pro
4	649	84.6	145	5	AAO19498 HSI prote
5	649	84.6	145	6	ABR58646 Human can
6	649	84.6	145	6	AD031800 Human nov
7	649	84.6	145	7	ADMA6959 Brain and
8	649	84.6	145	8	AD048475 Human PTH
9	621.5	81.0	180	7	ADMA6961 Brain and
10	472.5	61.6	149	7	ADMA6963 Brain and
11	329.5	43.0	92	5	AD081902 Human dio
12	294	38.3	54	8	AD048481 Mouse PTH
13	288	37.5	73	7	ADMA6962 Brain and
14	278	36.2	54	7	ADMA6960 Brain and
15	278	36.2	54	8	AD048477 Human PTH
16	272	35.5	80	7	ADMA6964 Brain and
17	140	18.3	25	7	ADMA6979 Brain and
18	99	12.9	18	7	ADMA6977 Brain and
19	89	11.6	16	8	AD048482 PTH respo
20	87.5	11.4	778	2	AAR13456 Dufy rec
21	84.5	11.0	684	4	AAB69330 Drosophi1
22	84	11.0	1433	5	ABP35624 Fungal ZB
23	83.5	10.9	742	5	ABR6153 Aspergill
24	82	10.7	574	5	ABG95028 Human tra
25	81.5	10.6	1100	4	AAG84930 Shrimp wh

ALIGNMENTS

26	80.5	10.5	600	7	ADJ70139	Adj70139 Human hea
27	80.5	10.5	643	8	ADI40859	Adi40859 Human kin
28	80.5	10.5	643	8	ABM83972	Abm83972 Human dia
29	80.5	10.5	852	4	AAM40296	Aam40296 Human pol
30	80.5	10.5	872	4	AAM40295	Aam40295 Human pol
31	80.5	10.5	886	4	AAM42081	Aam42081 Human pol
32	80.5	10.5	886	4	AAM42082	Aam42082 Human pol
33	80.5	10.5	950	8	AD089824	Adg89824 Antagonis
34	80.5	10.5	974	4	AADJ1506	Aau1506 Novel hum
35	80.5	10.5	2516	2	AAW01875	Aaw01875 Neuronal
36	80.5	10.5	2516	2	AAW01884	Aaw01884 Invertebr
37	80.5	10.5	2516	4	ABBE1076	Abb61076 Drosophi1
38	80.5	10.5	2516	5	ABJ17935	Abj17935 Drosophi1
39	80	10.4	516	3	AAU76758	Aau76758 Novel rec
40	80	10.4	1108	4	ABBS8571	Abbs8571 Drosophi1
41	80	10.4	1108	7	ADJ37991	Adj37991 D melanog
42	79.5	10.4	260	4	ABG09899	Abg09899 Novel hum
43	79.5	10.4	592	4	ABG15607	Abg15607 Novel hum
44	79	10.3	244	7	ABO69719	Abos69719 Pseudomon
45	78	10.2	724	2	AAI24750	Aay24750 Sarcophag

RESULT 1
AD048479 standard; protein; 145 AA.
AD048479; ADO48479;

12-AUG-2004 (first entry)

Mouse PTH responsive gene protein.

PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation; transgenic animal; osteopathic; gene therapy; osteoporosis; mouse;

murine.

Mus sp.

WO2004044152-A2.

27-MAY-2004.

10-NOV-2003; 2003WO-US035655.

12-NOV-2002; 2002US-0425532P.

(AMHP) WYETH.

Robinson JA, Stojanovic-Susulic V, Bablj P, Murrills RJ;

WPI; 2004-420299/39.

N-PSDB; AD048478.

New nucleic acid fragment encoding a PAIGB polypeptide, useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis.

Claim 9; SEQ ID NO 8; 169pp; English.

The invention relates to a novel PTH responsive gene (PAIGB) fragment encoding a polypeptide. The invention further comprises: a chimeric construct comprising the isolated nucleic acid fragment operatively linked to suitable regulatory sequences; a host cell transformed with the chimeric construct; a vector comprising the nucleic acid fragment; obtaining a nucleic acid fragment encoding the polypeptide; a method for obtaining a polypeptide; detecting the presence of the nucleic acid fragment; an antibody that specifically binds to one or more epitopes of a PAIGB polypeptide; a composition for regulating bone-forming activity in a mammal comprising the nucleic acid fragment, polypeptide or antibody; an agent that alters the expression of PAIGB gene or polypeptide;

CC determining whether an agent alters the expression of PAIGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.
XX

SQ Sequence 145 AA;

Query Match 100.0%; Score 767; DB 8; Length 145;

Best Local Similarity 100.0%; Pred. No. 2,8e-78;

Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAATDGPAGGLHAGVLEDEL 60
DB 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAATDGPAGGLHAGVLEDEL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQNLSSGPIITOKONGIATEAKRDARMSAR 120
DB 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQNLSSGPIITOKONGIATEAKRDARMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVAINTENIRQMDRSKRVTKNCIN 145

RESULT 2

ADO48473 standard; protein; 145 AA.

ADO48473;

12-AUG-2004 (first entry)

Rat PTH responsive gene protein.

XX PTH responsive gene; PAIGB; bone-forming; bone; bone density modulation;
KW transgenic animal; osteopathic; gene therapy; osteoporosis; rat.
XX

OS Rattus sp.

XX WO2004044152-A2.

PD 27-MAY-2004.

PF 10-NOV-2003; 2003WO-US035655.

PR 12-NOV-2002; 2002US-0425532P.

XX (AMHP) WYETH.

PI Robinson JA, Stojanovic-Suenilic V, Babi J, Murrills RJ;

XX WPI; 2004-420299/39.

DR N-PSDB; ADO48472.

XX New nucleic acid fragment encoding a PAIGB polypeptide, useful in
PT preparing a composition for diagnosing, treating or preventing bone
PT related disorders, e.g., osteoporosis.
XX
XX Claim 9, SEQ ID NO 2; 169pp; English.

XX The invention relates to a novel PTH responsive gene (PAIGB) fragment
CC encoding a polypeptide. The invention further comprises a chimeric
CC construct comprising the isolated nucleic acid fragment operatively
CC linked to suitable regulatory sequences, a host cell transformed with the
CC chimeric construct; a vector comprising the nucleic acid fragment;
CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
CC fragmenting a polypeptide; detecting the presence of the nucleic acid
CC fragment; an antibody that specifically binds to one or more epitopes of
CC a PAIGB polypeptide; a composition for regulating bone-forming activity
CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
CC ; an agent that alters the expression of PAIGB gene or polypeptide;
CC determining whether an agent alters the expression of PAIGB mRNA;
CC screening agents for effectiveness in altering expression of the nucleic
CC acid fragment; screening for agents useful for treating bone related
CC disorders; evaluating the efficacy of a treatment of a bone related
CC disorder in a subject; identifying polypeptides capable of binding to
CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
CC related agent; a transgenic animal comprising the DNA; an animal model
CC for the study of bone density modulation comprising a first group of
CC animals composed of the transgenic animal and a second group of control
CC animals; studying bone mass determinants; studying the modulation of bone
CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
CC for treating bone related disorders; identifying whether an agent which
CC has bone forming activity; and a stably transfected cell line comprising
CC two constructs, the first construct comprising a ligand binding domain
CC linked to a DNA binding domain which is linked to an activation domain
CC all of which expression is driven by a constitutive promoter, the second
CC construct comprising multiple copies of DNA binding elements linked to a
CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
CC polynucleotide has osteopathic activity. The PTH responsive gene may be
CC used to treat disorders by gene therapy. The nucleic acid is useful in
CC preparing a composition for diagnosing, treating or preventing bone
CC related disorders, e.g., osteoporosis. This sequence represents a PTH
CC responsive gene protein of the invention.
XX

SQ Sequence 145 AA;

Query Match 97.3%; Score 746; DB 8; Length 145;

Best Local Similarity 97.2%; Pred. No. 6,6e-76;

Matches 141; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAATDGPAGGLHAGVLEDEL 60
DB 1 MCGGSRADAIERPRYESTWRETETSTWLTYYTDSALPSAATDGPAGGLHAGVLEDEL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQNLSSGPIITOKONGIATEAKRDARMSAR 120
DB 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQNLSSGPIITOKONGIATEAKRDARMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVAINTENIRQMDRSKRVTKNCIN 145

RESULT 3

AAB95018 standard; protein; 145 AA.

AAB95018;

26-JUN-2001 (first entry)

XX Human protein sequence SEQ ID NO:16726.

KW Human; primer; detection; diagnosis; antisense therapy; gene therapy.

```

XX OS Homo sapiens.
XX EN EPI074617-A2.
XX PD 07-FEB-2001.
XX PF 26-JUL-2000; 2000EP-00116126.
XX PR 29-JUL-1999; 99JP-00248036.
XX PR 27-AUG-1999; 99JP-00300253.
XX PR 11-JAN-2000; 2000JP-00118776.
XX PR 02-MAY-2000; 2000JP-00183767.
XX PR 09-JUN-2000; 2000JP-00241899.
XX PA (HELI-) HELIX RES INST.
XX PI Oca T, Isogai T, Nishikawa T, Hayaishi K, Saito K, Yamamoto J,
XX PI Iehli S, Sugiyama T, Wakamatsu A, Nagai K, Otsuki T;
XX DR WPI; 2001-318749/34.
XX PT Primer sets for synthesizing polynucleotides, particularly the 5602 full-
XX PT length cDNAs defined in the specification, and for the detection and/or
XX PT diagnosis of the abnormality of the proteins encoded by the full-length
XX PT cDNAs.
XX PS Claim 8; SEQ ID NO 16726; 2537pp + Sequence Listing; English.
XX CC The present invention describes primer sets for synthesizing 5602 full-
XX CC length cDNAs defined in the specification. Where a primer set comprises:
XX CC (a) an oligo-dt primer and an oligonucleotide complementary to the
XX CC complementary strand of a polynucleotide which comprises one of the 5602
XX CC nucleotide sequences defined in the specification, where the
XX CC oligonucleotide comprises at least 15 nucleotides; or (b) a combination
XX CC of an oligonucleotide comprising a sequence complementary to the
XX CC complementary strand of a polynucleotide which comprises a 5'-end
XX CC sequence and an oligonucleotide comprising a sequence complementary to a
XX CC polynucleotide which comprises a 3'-end sequence; where the
XX CC oligonucleotide comprises at least 15 nucleotides and the combination of
XX CC the 5'-end sequence and 3'-end sequence is selected from those defined in the
XX CC specification. The primer sets can be used in antisense therapy and in
XX CC gene therapy. The primers are useful for synthesizing polynucleotides,
XX CC particularly full-length cDNAs. The primers are also useful for the
XX CC detection and/or diagnosis of the abnormality of the proteins encoded by
XX CC the full-length cDNAs. The primers allow obtaining of the full-length
XX CC cDNAs easily without any specialised methods. AAH03166 to AAH13628 and
XX CC AAH13633 to AAH18742 represent human cDNA sequences; AAB92446 to AAB95893
XX CC represent human amino acid sequences; and AAH13629 to AAH13632 represent
XX CC oligonucleotides, all of which are used in the exemplification of the
XX CC present invention
XX SQ Sequence 145 AA;
Query Match 84.6%; Score 649; DB 4; Length 145;
Best Local Similarity 83.4%; Pred. No. 6.3e-65;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
QY 1 MCGGSRADAIPRYESWTRETSTWLTYYTSDALPSAAATDSCGPAAGLHAGVLEDGL 60
DB 1 MCGGSRADAIPRYESWTRETSTWLTYYTSDALPSAAATDSCGPAAGLHAGVLEDGL 60
QY 61 SSGNGVLRPAAPGAIANPEKKMNCGTCCPNSONLSSGFLTKQNGLWATEAKDAXKMSAR 120
DB 61 PSNGVPRSTAPGAIANPEKKMNCGTCCPNSONLSSGFLTKQNGLWATEAKDAXKMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRTTKNCVN 145

```

RESULT 4
AAO19498

```

ID AAO19498 standard; protein; 145 AA.
XX AC AAO19498;
XX DT 20-DEC-2002 (first entry)
XX DE HSI protein variant.
XX DE HSI variant; cancer; tumour; unigene cluster; cytostatic; metastasis;
XX DE EST; expressed sequence tag; colon cancer; stomach cancer; breast cancer;
XX DE HSI69395; HSI27144; HSI2; HSI32793; HSI3.
XX OS Unidentified.
XX PN DE10103694-A1.
XX PD 01-AUG-2002.
XX PF 26-JAN-2001; 2001DE-01003694.
XX PR 26-JAN-2001; 2001DE-01003694.
XX PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.
XX PI Brett D, Kemmerer W;
XX PI WPI; 2002-644836/70.
XX DR N-PSDB; AAL50100.
XX PT Diagnosis and therapy of tumors, by determining expression rates of
XX PT specific expressed sequence tags of the unigene cluster, and subsequently
XX PT blocking their expression.
XX PS Claim 10; Page 5; 10pp; German.
XX CC The present invention relates to the use of expressed sequence tags
XX CC (ESTs), or variants, of the unigene cluster HSI69395 (HSI), HSI27144
XX CC (HS2) and/or HSI32793 (HS3) for diagnosis and therapy of tumors, in
XX CC which their expression rates in tumour cells and/or lymph nodes are
XX CC determined. The EST sequences are useful as prognostic markers of
XX CC survival of cancer patients (high levels of EST-related mRNA are
XX CC associated with a poor prognosis, specifically correlated with
XX CC development of metastases); and for diagnosis and/or therapy of solid
XX CC tumours, particularly of colon, stomach and breast. The present sequence
XX CC is a variant of the HSI protein shown in the exemplification of the
XX CC invention
XX SQ Sequence 145 AA;
Query Match 84.6%; Score 649; DB 5; Length 145;
Best Local Similarity 83.4%; Pred. No. 6.3e-65;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;
QY 1 MCGGSRADAIPRYESWTRETSTWLTYYTSDALPSAAATDSCGPAAGLHAGVLEDGL 60
DB 1 MCGGSRADAIPRYESWTRETSTWLTYYTSDALPSAAATDSCGPAAGLHAGVLEDGL 60
QY 61 SSGNGVLRPAAPGAIANPEKKMNCGTCCPNSONLSSGFLTKQNGLWATEAKDAXKMSAR 120
DB 61 PSNGVPRSTAPGAIANPEKKMNCGTCCPNSONLSSGFLTKQNGLWATEAKDAXKMSAR 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRTTKNCVN 145

```

RESULT 5
ABR58646
ID ABR58646 standard; protein; 145 AA.
XX AC ABR58646;
XX DT 09-JUL-2003 (first entry)

XX DE Human cancer related protein SEQ ID NO:303.
 XX XX
 XX KW Human; cancer; diagnosis; screening; modulator; leukemia; ischaemia;
 XX KW heart disease; atherosclerosis; endometriosis.
 XX OS Homo sapiens.
 XX PN WO2003025138-A2.
 XX XX
 XX PD 27-MAR-2003.
 XX XX
 XX PF 17-SEP-2002; 2002WO-US029560.
 XX PR 17-SEP-2001; 2001US-0323469P.
 XX PR 20-SEP-2001; 2001US-0323887P.
 XX PR 13-NOV-2001; 2001US-0350666P.
 XX PR 08-FEB-2002; 2002US-0355145P.
 XX PR 08-FEB-2002; 2002US-0355257P.
 XX PR 12-APR-2002; 2002US-0372246P.
 XX XX
 XX PA (EOSB-) EOS BIOTECHNOLOGY INC.
 XX XX
 XX PI Afar D, Aziz N, Gish KC, Hevezi PA, Mack DH, Wilson KE;
 XX PI Zlotnick A;
 XX XX
 XX DR WPI: 2003-354600/33.
 XX DR N-PSDB; ACC72796.
 XX XX
 XX PT New genes that are up-regulated or down-regulated in cancers, useful as
 XX PT markers for diagnosing e.g. cancer, ischemia or heart diseases, or as
 XX PT therapeutic targets for screening drugs for treating these diseases.
 XX PS
 XX Claim 12; Page 753; 767pp; English.
 XX XX
 XX CC The present invention describes an isolated nucleic acid molecule, which
 XX CC comprises the sequence of any of the genes that are up-regulated or down-
 XX CC regulated in specific cancers (e.g. about 1031 genes up-regulated in
 XX CC acute lymphocytic leukemia). ACC72641 to ACC72860 represent cancer
 XX CC related gene nucleotide sequences which encode the proteins given in
 XX CC ABR8521 to ABR85709. Also described: (1) determining the presence or
 XX CC absence of a pathological cell in a patient; (2) an expression vector
 XX CC comprising a nucleic acid molecule described above; (3) a host cell
 XX CC comprising the vector; (4) an isolated polypeptide, which is encoded by
 XX CC the nucleic acid; (5) an antibody that specifically binds the polypeptide
 XX CC of (4); (6) specifically targeting a compound to a pathological cell in a
 XX CC patient by administering to the patient the antibody above; and (7) a
 XX CC drug screening assay. The nucleic acid is useful as diagnostic markers or
 XX CC therapeutic targets. In particular, the nucleic acid is useful for
 XX CC diagnosing a pathology, e.g. cancer (e.g. cancer of the bone marrow,
 XX CC bladder, brain, breast, cervix, colon/rectum, kidney, lung, ovary,
 XX CC pancreas, prostate, skin and uterus), wounds, ischemia, heart diseases,
 XX CC atherosclerosis and endometriosis. The nucleic acid is also useful in
 XX CC drug screening, particularly for identifying agents for treating these
 XX CC pathologies
 XX XX
 XX SQ Sequence 145 AA;
 XX XX
 XX Query Match 84.6%; Score 649; DB 6; Length 145;
 XX Best Local Similarity 83.4%; Pred. No. 6,3e-65;
 XX Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

RESULT 6
 ADCC31800
 ID ADCC31800 standard; protein; 145 AA.
 XX XX
 XX AC ADCC31800;
 XX XX
 XX DT 18-DEC-2003 (first entry)
 XX XX
 XX DE Human novel polypeptide sequence, SEQ ID NO:1882.
 XX XX
 XX KW Human; diagnostic; drug screening; forensics; gene mapping;
 XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 XX KW ulcers; osteoporosis; autoimmune disease; cancer;
 XX KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 XX KW neuroprotective; anti-nausea; anticoagulant; thrombolytic; vularary;
 XX KW antulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 XX KW gene therapy; chromosome 8.
 XX XX
 XX OS Homo sapiens.
 XX XX
 XX PN WO2003029271-A2.
 XX XX
 XX PD 10-APR-2003.
 XX XX
 XX PF 24-SEP-2002; 2002WO-US030474.
 XX XX
 XX PR 24-SEP-2001; 2001US-0324631P.
 XX XX
 XX PA (HYSE-) HYSEQ INC.
 XX XX
 XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 XX PI Haley-Vicente D, Dymanc RT;
 XX XX
 XX DR WPI: 2003-371981/35.
 XX DR N-PSDB; ADCC30829.
 XX XX
 XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 XX PT treating conditions such as neurodegenerative diseases, anemias, platelet
 XX PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 XX PT cancer.
 XX PS
 XX Claim 20; SEQ ID NO 1882; 1185pp; English.
 XX XX
 XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 XX CC ADC30889) and the polypeptides they encode (ADCC30890-ADCC31860). The
 XX CC invention also relates to nucleic acid sequences over 99% identical with
 XX CC the novel human cDNAs. The invention additionally encompasses expression
 XX CC vectors and host cells comprising a nucleic acid of the invention; the
 XX CC recombinant production of a polypeptide of the invention; an antibody
 XX CC against a polypeptide of the invention; a method of detecting
 XX CC polynucleotides or polypeptides of the invention; and methods of
 XX CC identifying a compound which binds to a polypeptide of the invention. The
 XX CC invention further discloses methods of preventing, treating or
 XX CC ameliorating a medical condition; kits comprising polynucleotide probes
 XX CC and/or monoclonal antibodies for carrying out the methods of the
 XX CC invention; methods for the identification of compounds that modulate the
 XX CC expression or activity of the polynucleotide and/or polypeptide; and 767
 XX CC contig sequences corresponding to the cDNA sequences of the invention
 XX CC (ADCC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 XX CC -ADCC3394). The nucleic acids and polypeptides of the invention are
 XX CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 XX CC identification of mutations responsible for genetic disorders or other
 XX CC traits, for assessing biodiversity, and in producing many other types of
 XX CC data and products dependent on DNA and amino acid sequences. They are
 XX CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 XX CC disease and other neurodegenerative diseases, anaemia, platelet
 XX CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 XX CC cancer. The nucleic acids may also be used as hybridisation probes or
 XX CC primers, and in the recombinant production of a protein. The polypeptides

are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human polypeptide sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 145 AA:

Query Match 84.6%; Score 649; DB 7; Length 145;
Best Local Similarity 83.4%; Pred. No. 6,3e-65;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSGRADAIERYESWTRTESTWLTYYTSDALPSAAATDSGPBAGLHAGVLEDL 60
DB 1 MCGGSGRADAIERYESWTRTESTWLTYYTSDAPSAAPDSGPBAGLHAGVLEDL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQNSLSSGPLTKONGLMATEAKDKAKMSAR 120
DB 61 PSNGVPRSTAPGGIPNPEKKNCECTCCPNQNSLSSGPLTKONGLOTTEAKDKAKMPAK 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 7

ADM46959 standard; protein; 145 AA.

ADM46959;

03-JUN-2004 (first entry)

Brain and Acute Leukemia, Cytoplasmic alternate protein #1.

acute myelogenous leukemia; gene expression; BAAIC;

chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;

Cytoplasmic; exon.

Homo sapiens.

Key Location/Qualifiers

Misc-difference 41

W02003040347-A2.

15-MAY-2003.

12-NOV-2002; 2002MO-US036375.

09-NOV-2001; 2001US-0348210P.

(OHIS) UNIV OHIO STATE RES FOUND.

Tanner SM, De La Chapell A;

WPI; 2003-441564/41.

N-PSDB; ADM46951.

Characterizing acute or chronic myelogenous leukemia, or prostate cancer in a patient comprising assaying for the overexpression of one or more BAAIC transcripts in cells obtained from the patient.

Disclosure; SEQ ID NO 17; 78bp; English.

The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAAIC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful

for detecting BAAIC overexpression. This sequence corresponds to a BAAIC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

Sequence 145 AA:

Query Match 84.6%; Score 649; DB 7; Length 145;
Best Local Similarity 83.4%; Pred. No. 6,3e-65;
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSGRADAIERYESWTRTESTWLTYYTSDALPSAAATDSGPBAGLHAGVLEDL 60
DB 1 MCGGSGRADAIERYESWTRTESTWLTYYTSDAPSAAPDSGPBAGLHAGVLEDL 60
QY 61 SSGVLRPAAPGGIANPEKKNCGTCCPNQNSLSSGPLTKONGLMATEAKDKAKMSAR 120
DB 61 PSNGVPRSTAPGGIPNPEKKNCECTCCPNQNSLSSGPLTKONGLOTTEAKDKAKMPAK 120
QY 121 EVAINTENIRQMDRSKRVTKNCIN 145
DB 121 EVTINVTDSIQMDRSRRITKNCVN 145

RESULT 8

ADO48475 standard; protein; 145 AA.

ADO48475;

12-AUG-2004 (first entry)

Human PTH responsive gene protein.

PTH responsive gene; PARG, bone-forming; bone, bone density modulation;

transgenic animal; osteopathic; gene therapy; osteoporosis; human.

Homo sapiens.

W02004044152-A2.

27-MAY-2004.

10-NOV-2003; 2003MO-US035655.

12-NOV-2002; 2002US-0425532P.

(AMHP) WYETH.

Robinson JA, Stojanovic-Susulic V, Babij P, Murrill RJ;

WPI; 2004-420299/39.

N-PSDB; ADO48474.

New nucleic acid fragment encoding a PARG polypeptide, useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis.

Claim 9; SEQ ID NO 4; 169bp; English.

The invention relates to a novel PTH responsive gene (PARG) fragment encoding a polypeptide. The invention further comprises a chimeric construct comprising the isolated nucleic acid fragment operatively linked to suitable regulatory sequences; a host cell transformed with the chimeric construct; a vector comprising the nucleic acid fragment; obtaining a polypeptide; detecting the presence of the nucleic acid fragment; an antibody that specifically binds to one or more epitopes of a PARG polypeptide; a composition for regulating bone-forming activity in a mammal comprising the nucleic acid fragment, polypeptide or antibody; an agent that alters the expression of PARG gene or polypeptide; determining whether an agent alters the expression of PARG mRNA; screening agents for effectiveness in altering expression of the nucleic acid fragment; screening for agents useful for treating bone related

CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PAIGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAIGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain
 CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAIGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAIGB gene is induced. The PAIGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosis, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.

CC Sequence 145 AA;

Query Match 84.6%; Score 649; DB 8; Length 145;
 Best Local Similarity 83.4%; Pred. No. 6.3e-65;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAAATDSGPEAGGLHAGVLEDEL 60
 DB 1 MCGGSRADAIERYESWTRETESTWLTYYTSDAPPSAAAPDSGPEAGGLHSGMLEDEL 60
 QY 61 SNGVLRPAAPGGIAPPEKMNCGTCNPSONLSGGELTGOKGWLATKAKDKAKMSAR 120
 DB 61 PSNGVLRSTAPGCIPEPEKKNCECTCPNPQSLSSGELTGOKGLOTTEAKDKAKMPAK 120
 QY 121 EVAINTENIRQMDRSKRYTKNCIN 145
 DB 121 EVTINTDSIQQMDRSRRITKNCVN 145

RESULT 9
 ADM46961
 ID ADM46961 standard; protein; 180 AA.

AC ADM46961;

DT 03-JUN-2004 (first entry)

DB Brain and Acute Leukemia, Cytoplasmic alternate protein #3.

KW acute myelogenous leukemia; gene expression; BAALC;
 KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 KW Cytoplasmic; exon.

OS Homo sapiens.

Key Location/Qualifiers
 FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002MO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapel1 A;

XX

DR WPI; 2003-441564/41.
 DR N-PSDB; ADM46953.

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprises assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 19; 78bp; English.

XX The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.

CC Sequence 180 AA;

Query Match 81.0%; Score 621.5; DB 7; Length 180;
 Best Local Similarity 67.2%; Pred. No. 1.1e-61;
 Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCGGSRADAIERYESWTRETESTWLTYYTSDALPSAAATDSGPEAGGLHAGVLEDEL 53
 DB 1 MCGGSRADAIERYESWTRETESTWLTYYTSDAPPSAAAPDSGPEAGGLHSGVLEAKS 60
 QY 54 -----GVLEGLSSNGVLRPAAPGGIAPPEKMNCGT 85
 DB 61 KIKAPLTVDSDGLPSASKMAPLAVFSGHMLEDGLPSNGVLRSTAPGCIPEPEKKNCECT 120
 QY 86 QCPNQNLSGGELTGOKGWLATKAKMSAREVAINVENTENIRQMDRSKRYTKNCIN 145
 DB 121 QCPNPQSLSSGELTGOKGLOTTEAKDKAKMPAKEVTINTDSIQQMDRSRRITKNCVN 180

RESULT 10

ADM46963
 ID ADM46963 standard; protein; 149 AA.

AC ADM46963;

DT 03-JUN-2004 (first entry)

DB Brain and Acute Leukemia, Cytoplasmic alternate protein #5.

KW acute myelogenous leukemia; gene expression; BAALC;
 KW chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia;
 KW Cytoplasmic; exon.

OS Homo sapiens.

Key Location/Qualifiers
 FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

PD 15-MAY-2003.

PF 12-NOV-2002; 2002MO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

PA (OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De La Chapel1 A;

WPI; 2003-441564/41.

DR N-PSDB; ADM46955.

XX

PT Characterizing acute or chronic myelogenous leukemia, or prostate cancer
 PT in a patient comprising assaying for the overexpression of one or more
 PT BAALC transcripts in cells obtained from the patient.
 XX
 PS Disclosure; SEQ ID NO 21; 78pp; English.
 XX
 CC The invention relates to a method of characterizing acute myelogenous
 CC leukemia (AML) in a patient by assaying for the overexpression of one or
 CC more BAALC transcripts in cells obtained from the patient, where an
 CC overexpression indicates that the patient has an aggressive form of AML.
 CC The methods, kits and probes are useful for characterizing acute or
 CC chronic myelogenous leukemia, or prostate cancer. They are also useful
 CC for detecting BAALC overexpression. This sequence corresponds to a BAALC
 CC (brain and acute leukemia, Cytoplasmic) protein encoded by an alternative
 CC spliced RNA consisting of exons 1, 6 and 8.
 XX
 SQ Sequence 149 AA;
 Query Match 61.6%; Score 472.5; DB 7; Length 149;
 Best Local Similarity 65.3%; Pred. No. 6,4e-45;
 Matches 94; Conservative 3; Mismatches 12; Indels 35; Gaps 1;
 QY 1 MCGCGSRADAIEPRYYSWTRETESTWLTYSDDALPSAATDSGPAGGLHA----- 53
 1 MCGCGSRADAIEPRYYSWTRETESTWLTYSDDALPSAATDSGPAGGLHSLAEKS 60
 DB 54 -----GVLDEGLSNGVLRPAAPGCIANPEKKNCGT 85
 61 KIKAPTDVSDEGLPSASKAPLAVFSGHMLEDLPSNGVPRSTAPGGINPEKKNCT 120
 QY 86 QCPNQLSSGGLTKOKNGLWATE 109
 121 QCPNQLSSGGLTKOKNGLOTTE 144
 DB
 RESULT 11
 ADQ81902
 ID ADQ81902 standard; protein; 92 AA.
 AC ADQ81902;
 XX
 DT 09-SEP-2004 (first entry)
 XX
 DE Human dioxigenase 10.12.
 XX
 KW Human; enzyme; dioxigenase 10.12; malignant tumour; inflammation;
 KW Immunological disease; haemopathy; HIV infection.
 XX
 OS Homo sapiens.
 XX
 PN CN1344798-A.
 XX
 PD 17-APR-2002.
 XX
 PF 29-SEP-2000; 2000CN-00125495.
 XX
 PR 29-SEP-2000; 2000CN-00125495.
 XX
 PA (SHAN-) SHANGHAI BIODOOR GENE DEV CO LTD.
 XX
 PI Mao Y, Xie Y;
 XX
 DR WPI, 2002-509506/55.
 XX
 DR N-PSDB; ADQ81901.
 XX
 PT New polypeptide human dioxigenase 10.12 and polynucleotides encoding this
 PT polypeptide, useful for treating various diseases, such as malignant
 PT tumors, inflammations, immunological diseases, hemopathy and HIV
 PT infection.
 XX
 PS Claim 1; SEQ ID NO 2; 33pp; Chinese.
 XX
 CC The present invention discloses a new kind of polypeptide, human

CC dioxigenase 10.12, polynucleotides encoding this polypeptide, a DNA
 CC recombination process to produce the polypeptide and antagonist against
 CC the polypeptide. The present invention also discloses the method of
 CC applying the polypeptide in treating various diseases, such as malignant
 CC tumors, inflammations, immunological diseases, haemopathy and HIV
 CC infection. The present sequence is the human dioxigenase 10.12.
 XX
 SQ Sequence 92 AA;
 Query Match 43.0%; Score 329.5; DB 5; Length 92;
 Best Local Similarity 63.4%; Pred. No. 5e-29;
 Matches 64; Conservative 8; Mismatches 10; Indels 19; Gaps 1;
 QY 45 GPBAGGLHAGVLEDGLSNGVLRPAAPGCIANPEKKNCGTQCPNQLSSGGLTKOKNG 104
 11 GPBAGGLHAGVLEDGLSNGVLRPAAPGCIANPEKKNCGTQCPNQLSSGGLTKOKNG 51
 DB 105 LMAFEAKRDAKRMASREVAINTENTROMDRSKRTKNCIN 145
 52 LQTFEAKRDAKRMASREVAINTENTROMDRSKRTKNCIN 92
 QY
 DB
 RESULT 12
 ADQ48481
 ID ADQ48481 standard; protein; 54 AA.
 AC ADQ48481;
 XX
 DT 12-AUG-2004 (first entry)
 XX
 DE Mouse PTH responsive gene protein exon 2 splice variant.
 XX
 KW PTH responsive gene; PARGB, bone-forming; bone; bone density modulation;
 KW transgenic animal; osteopathic; gene therapy; osteoporosis; mouse;
 XX
 OS Mus sp.
 XX
 PN WO2004044152-A2.
 XX
 PD 27-MAY-2004.
 XX
 PF 10-NOV-2003; 2003WO-US035655.
 XX
 PR 12-NOV-2002; 2002US-0425532P.
 XX
 PA (AMHP) WYETH.
 XX
 PI Robinson JA, Stojanovic-Susulic V, Babi P, Murrills RJ;
 XX
 DR WPI, 2004-420299/39.
 XX
 DR N-PSDB; ADQ48480.
 XX
 PT New nucleic acid fragment encoding a PARGB polypeptide, useful in
 PT preparing a composition for diagnosing, treating or preventing bone
 PT related disorders, e.g., osteoporosis.
 XX
 PS Claim 9; SEQ ID NO 10; 169pp; English.
 XX
 CC The invention relates to a novel PTH responsive gene (PARGB) fragment
 CC encoding a polypeptide. The invention further comprises: a chimeric
 CC construct comprising the isolated nucleic acid fragment operatively
 CC linked to suitable regulatory sequences; a host cell transformed with the
 CC chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 CC fragment; an antibody that specifically binds to one or more epitopes of
 CC a PARGB polypeptide; a composition for regulating bone-forming activity
 CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 CC; an agent that alters the expression of PARGB gene or polypeptide;
 CC determining whether an agent alters the expression of PARGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related

disorders; evaluating the efficacy of a treatment of a bone related disorder in a subject; identifying polypeptides capable of binding to PAIGB; monitoring the effectiveness of treatment of a subject with a bone related agent; a transgenic animal comprising the DNA; an animal model for the study of bone density modulation comprising a first group of animals composed of the transgenic animal and a second group of control animals; studying bone mass determinants; studying the modulation of bone mass; studying an effect of PAIGB on bone disorders; identifying an agent for treating bone related disorders; identifying whether an agent which has bone forming activity; and a stably transfected cell line comprising two constructs, the first construct comprising a ligand binding domain linked to a DNA binding domain which is linked to an activation domain all of which expression is driven by a constitutive promoter, the second construct comprising multiple copies of DNA binding elements linked to a minimal promoter which is linked to PAIGB cDNA, where upon the addition of chemical inducer, transcription of PAIGB gene is induced. The PAIGB polynucleotide has osteopathic activity. The PTH responsive gene may be used to treat disorders by gene therapy. The nucleic acid is useful in preparing a composition for diagnosing, treating or preventing bone related disorders, e.g., osteoporosis. This sequence represents a PTH responsive gene protein of the invention.

SO Sequence 54 AA;

Query Match 38.3%; Score 294; DB 8; Length 54;
Best Local Similarity 100.0%; Pred. No. 2,4e-25;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERRYESWTRTESTWLTYSDDALPSAAATDSGPEAGGHLHAG 54
DB 1 MCGGSRADAIERRYESWTRTESTWLTYSDDALPSAAATDSGPEAGGHLHAG 54

RESULT 13

ADM46962
ID ADM46962 standard; protein; 73 AA.

AC ADM46962;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #4.

KW acute myelogenous leukemia; gene expression; BAALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia; Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

XX 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

(OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De la Chapell A;

DR WPI; 2003-441564/41.

XX N-PSDB; ADM46954.

Characterizing acute or chronic myelogenous leukemia, or prostate cancer in a patient comprises assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 20; 78pp; English.

XX The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful for detecting BAALC overexpression. This sequence corresponds to a BAALC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 73 AA;

Query Match 37.5%; Score 288; DB 7; Length 73;
Best Local Similarity 91.4%; Pred. No. 1,8e-24;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERRYESWTRTESTWLTYSDDALPSAAATDSGPEAGGHLHAG 58
DB 1 MCGGSRADAIERRYESWTRTESTWLTYSDDALPSAAATDSGPEAGGHLHAG 58

RESULT 14

ADM46960
ID ADM46960 standard; protein; 54 AA.

AC ADM46960;

DT 03-JUN-2004 (first entry)

DE Brain and Acute Leukemia, Cytoplasmic alternate protein #2.

KW acute myelogenous leukemia; gene expression; BAALC;

KM chronic myelogenous leukemia; prostate cancer; Brain and Acute Leukemia; Cytoplasmic; exon.

OS Homo sapiens.

FH Key Location/Qualifiers

FT Misc-difference 41 /note= "encoded by GCS"

PN WO2003040347-A2.

XX 15-MAY-2003.

PF 12-NOV-2002; 2002WO-US036375.

PR 09-NOV-2001; 2001US-0348210P.

(OHIS) UNIV OHIO STATE RES FOUND.

PI Tanner SM, De la Chapell A;

DR WPI; 2003-441564/41.

XX N-PSDB; ADM46952.

Characterizing acute or chronic myelogenous leukemia, or prostate cancer in a patient comprises assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient.

PS Disclosure; SEQ ID NO 18; 78pp; English.

XX The invention relates to a method of characterizing acute myelogenous leukemia (AML) in a patient by assaying for the overexpression of one or more BAALC transcripts in cells obtained from the patient, where an overexpression indicates that the patient has an aggressive form of AML. The methods, kits and probes are useful for characterizing acute or chronic myelogenous leukemia, or prostate cancer. They are also useful for detecting BAALC overexpression. This sequence corresponds to a BAALC (Brain and Acute Leukemia, Cytoplasmic) protein encoded by an alternative spliced RNA consisting of exons 1, 6 and 8.

SQ Sequence 54 AA;

Query Match 36.2%; Score 278; DB 7; Length 54;
 Best Local Similarity 94.4%; Pred. No. 1.6e-23;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MCGGSRADAIEPRYESTWRETSTWLTITDSDALPSAATDGPAGGLHAG 54
 |||||
 Db 1 MCGGSRADAIEPRYESTWRETSTWLTITDSDALPSAATDGPAGGLHAG 54

RESULT 15
 ADO48477
 ID ADO48477 standard; protein; 54 AA.
 AC ADO48477;
 DT 12-AUG-2004 (first entry)
 XX Human PTH responsive gene protein exon 2 splice variant.
 DE
 XX PTH responsive gene; PAlGB, bone-forming; bone; bone density modulation;
 KW transgenic animal; osteopathic; gene therapy; osteoporosis; human.
 XX
 OS Homo sapiens.
 XX MO200404152-A2.
 PN
 XX 27-MAY-2004.
 PD
 XX 10-NOV-2003; 2003WO-US035655.
 PF
 XX 12-NOV-2002; 2002US-0425532P.
 PR
 XX (AMHP) WYETH.
 PA
 XX Robinson JA, Stojanovic-Susulic V, Babić P, Murrills RJ;
 PI WPI, 2004-420299/39.
 DR N-PSDB; ADO48476.
 XX
 PT New nucleic acid fragment encoding a PAlGB polypeptide, useful in
 PT preparing a composition for diagnosing, treating or preventing bone
 PT related disorders, e.g., osteoporosis.
 PS Claim 9; SEQ ID NO 6; 169bp; English.
 XX
 CC The invention relates to a novel PTH responsive gene (PAlGB) fragment
 CC encoding a polypeptide. The invention further comprises: a chimeric
 CC construct comprising the isolated nucleic acid fragment operatively
 CC linked to suitable regulatory sequences; a host cell transformed with the
 CC chimeric construct; a vector comprising the nucleic acid fragment;
 CC obtaining a nucleic acid fragment encoding the polypeptide; a method for
 CC obtaining a polypeptide; detecting the presence of the nucleic acid
 CC fragment; an antibody that specifically binds to one or more epitopes of
 CC a PAlGB polypeptide; a composition for regulating bone-forming activity
 CC in a mammal comprising the nucleic acid fragment, polypeptide or antibody
 CC; an agent that alters the expression of PAlGB gene or polypeptide;
 CC determining whether an agent alters the expression of PAlGB mRNA;
 CC screening agents for effectiveness in altering expression of the nucleic
 CC acid fragment; screening for agents useful for treating bone related
 CC disorders; evaluating the efficacy of a treatment of a bone related
 CC disorder in a subject; identifying polypeptides capable of binding to
 CC PAlGB; monitoring the effectiveness of treatment of a subject with a bone
 CC related agent; a transgenic animal comprising the DNA; an animal model
 CC for the study of bone density modulation comprising a first group of
 CC animals composed of the transgenic animal and a second group of control
 CC animals; studying bone mass determinants; studying the modulation of bone
 CC mass; studying an effect of PAlGB on bone disorders; identifying an agent
 CC for treating bone related disorders; identifying whether an agent which
 CC has bone forming activity; and a stably transfected cell line comprising
 CC two constructs, the first construct comprising a ligand binding domain
 CC linked to a DNA binding domain which is linked to an activation domain

CC all of which expression is driven by a constitutive promoter, the second
 CC construct comprising multiple copies of DNA binding elements linked to a
 CC minimal promoter which is linked to PAlGB cDNA, where upon the addition
 CC of chemical inducer, transcription of PAlGB gene is induced. The PAlGB
 CC polynucleotide has osteopathic activity. The PTH responsive gene may be
 CC used to treat disorders by gene therapy. The nucleic acid is useful in
 CC preparing a composition for diagnosing, treating or preventing bone
 CC related disorders, e.g., osteoporosis. This sequence represents a PTH
 CC responsive gene protein of the invention.

SQ Sequence 54 AA;

Query Match 36.2%; Score 278; DB 8; Length 54;
 Best Local Similarity 94.4%; Pred. No. 1.6e-23;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 MCGGSRADAIEPRYESTWRETSTWLTITDSDALPSAATDGPAGGLHAG 54
 |||||
 Db 1 MCGGSRADAIEPRYESTWRETSTWLTITDSDALPSAATDGPAGGLHAG 54

Search completed: April 12, 2005, 15:37:19
 Job time : 104.667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:28:13 ; Search time 26.6667 Seconds
(without alignments)
405.904 Million cell updates/sec

Title: US-10-705-716a-8

Perfect score: 767
Sequence: 1 MCGGSRADAIERPRYESWT.....VTENIRQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgnt2_6/prodata/1/1aa/5A COMB pep: *
2: /cgnt2_6/prodata/1/1aa/5B COMB pep: *
3: /cgnt2_6/prodata/1/1aa/6A COMB pep: *
4: /cgnt2_6/prodata/1/1aa/6B COMB pep: *
5: /cgnt2_6/prodata/1/1aa/6CTUS COMB pep: *
6: /cgnt2_6/prodata/1/1aa/backfile1 pep: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87.5	11.4	778	6 5198347-4	Patent No. 5198347
2	87.5	11.4	778	6 5198347-4	Patent No. 5198347
3	80.5	10.5	2516	3 US-08-374-077C-2	Sequence 2, App11
4	80.5	10.5	2516	3 US-08-895-590-2	Sequence 2, App11
5	80.5	10.5	2516	4 US-09-539-879A-2	Sequence 2, App11
6	79.5	10.4	694	4 US-09-949-016-7117	Sequence 7117, Ap
7	79	10.3	244	4 US-09-252-991A-18465	Sequence 18465, A
8	78	10.2	724	3 US-09-121-964-1	Sequence 1, App11
9	77	10.0	715	4 US-09-252-991A-32740	Sequence 32740, A
10	75.5	9.8	286	4 US-09-071-035-176	Sequence 176, App
11	75.5	9.8	305	4 US-09-071-035-174	Sequence 174, App
12	75.5	9.8	700	3 US-09-408-647A-2	Sequence 2, App11
13	74	9.6	726	4 US-09-328-352-8115	Sequence 8115, Ap
14	74	9.6	1113	4 US-09-618-425-9	Sequence 9, App11
15	73.5	9.6	346	4 US-09-538-092-723	Sequence 723, App
16	73	9.5	316	4 US-09-248-796A-16753	Sequence 16753, A
17	73	9.5	512	4 US-09-496-320-11	Sequence 11, App1
18	72.5	9.5	1050	4 US-09-538-092-131	Sequence 131, App
19	72	9.4	386	4 US-09-328-352-7679	Sequence 7679, Ap
20	72	9.4	853	3 US-09-254-325-2	Sequence 2, App11
21	72	9.4	866	3 US-09-079-415-6	Sequence 6, App11
22	72	9.4	866	3 US-08-750-458A-2	Sequence 2, App11
23	71.5	9.3	240	4 US-09-252-991A-21991	Sequence 21991, A
24	71.5	9.3	258	4 US-09-270-767-43579	Sequence 43579, A
25	71.5	9.3	384	4 US-09-270-767-42538	Sequence 42538, A
26	71.5	9.3	412	4 US-09-949-016-6666	Sequence 6666, Ap
27	71.5	9.3	415	4 US-09-949-016-11105	Sequence 11105, A

28	71.5	9.3	415	4 US-09-949-016-11106	Sequence 11106, A
29	71.5	9.3	3623	4 US-09-341-461-2	Sequence 2, App11
30	71	9.3	224	4 US-09-270-767-59848	Sequence 59848, A
31	71	9.3	433	4 US-09-270-767-44417	Sequence 44417, A
32	71	9.3	754	4 US-09-252-991A-18938	Sequence 18938, A
33	70.5	9.2	172	4 US-09-252-991A-17765	Sequence 17765, A
34	70.5	9.2	226	4 US-09-252-991A-30546	Sequence 30546, A
35	70.5	9.2	351	4 US-09-585-645A-11	Sequence 11, App1
36	70.5	9.2	351	4 US-09-585-645A-46	Sequence 46, App1
37	70	9.1	217	4 US-09-252-991A-26962	Sequence 26962, A
38	70	9.1	266	4 US-09-252-991A-27673	Sequence 27673, A
39	70	9.1	360	3 US-09-509-902A-9	Sequence 9, App11
40	70	9.1	546	4 US-09-235-153-3	Sequence 3, App11
41	70	9.1	662	4 US-09-543-681A-7467	Sequence 7467, Ap
42	69.5	9.1	139	4 US-09-252-991A-19245	Sequence 19245, A
43	69.5	9.1	219	4 US-09-270-767-57118	Sequence 57118, A
44	69.5	9.1	323	4 US-09-270-767-41874	Sequence 41874, A
45	69	9.0	245	4 US-09-252-991A-19505	Sequence 19505, A

ALIGNMENTS

```
RESULT 1
5198347-4
; Patent No. 5198347
; APPLICANT: MILLER, LOUIS H.; ADAMS, JOHN H.; KASLOW,
; DAVIC C.; FANG, XIANGDONG
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
; PLASMODIUM KNOWLEST DUFFY RECEPTOR
; NUMBER OF SEQUENCES: 27
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,837
; FILING DATE: 20-JUL-1990
; SEQ ID NO:4:
; LENGTH: 778
5198347-4
; Patent No. 5198347
; APPLICANT: MILLER, LOUIS H.; ADAMS, JOHN H.; KASLOW,
; DAVIC C.; FANG, XIANGDONG
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
; PLASMODIUM KNOWLEST DUFFY RECEPTOR
; NUMBER OF SEQUENCES: 27
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,837
; FILING DATE: 20-JUL-1990
; SEQ ID NO:4:
; LENGTH: 778
5198347-4

Query Match      11.4%; Score 87.5; DB 6; Length 778;
Best Local Similarity 31.8%; Pred. No. 0.44;
Matches 28; Conservative 13; Mismatches 44; Indels 3; Gaps 3;

QY      31 TSDALPSAARDSCGEAGAGVLEDTLSNGV-LRPAAPGAIANPEKKNC-GTQCP 88
DB      352 TVSSDVPVSGKDSGSTSASHALAGENGVEVNGTDEKEDKADPKDIEVKKQDT 411

QY      89 NSQNLSS-GPLTQKONGLMATEAKRDAK 115
DB      412 DDRSGSLGPHTDERATLGETMEKOTE 439

RESULT 2
5198347-4
; Patent No. 5198347
; APPLICANT: MILLER, LOUIS H.; ADAMS, JOHN H.; KASLOW,
; DAVIC C.; FANG, XIANGDONG
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
; PLASMODIUM KNOWLEST DUFFY RECEPTOR
; NUMBER OF SEQUENCES: 27
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,837
; FILING DATE: 20-JUL-1990
; SEQ ID NO:4:
; LENGTH: 778
5198347-4

Query Match      11.4%; Score 87.5; DB 6; Length 778;
Best Local Similarity 31.8%; Pred. No. 0.44;
Matches 28; Conservative 13; Mismatches 44; Indels 3; Gaps 3;

QY      31 TSDALPSAARDSCGEAGAGVLEDTLSNGV-LRPAAPGAIANPEKKNC-GTQCP 88
DB      352 TVSSDVPVSGKDSGSTSASHALAGENGVEVNGTDEKEDKADPKDIEVKKQDT 411
```

QY 89 NSQNLSS-GPLTQKONGLWATEAKRDAK 115
Db 412 DPRSQGLSPHTDERATLGETHMKDTE 439

RESULT 3
US-08-374-077C-2
; Sequence 2, Application US/08374077C
; Patent No. 6027912
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha
; TITLE OF INVENTION: Calcium Channel Subunit
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22314-3187
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/374,077C
; FILING DATE: 19-JAN-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm M.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 022650-264
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-836-6620
; TELEFAX: 703-836-2021
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2516 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-374-077C-2

Query Match 10.5%; Score 80.5; DB 3; Length 2516;
Best Local Similarity 22.4%; Pred. No. 15;
Matches 39; Conservative 23; Mismatches 59; Indels 53; Gaps 8;
QY 3 CCGSRADAIEPRY--YESWTRETESTWLTYYDSALPSAATDSCPEAGLHAGVLEBGL 60
Db 207 CCGGGISAPPRRLTPEBAWLOQPQ-----NSVTSAGSTNSFSSG-----GR 249
QY 61 SSGVLRPAAPGCIANPEKKNK-----GTQC-----PNS 90
Db 250 DDNSGY--SAVGGDSSSSNSCNCIDITGDNSTLHGLGVGVCSFIADCDNSEDGDDPNN 307
QY 91 QNLSSGPL-TQKONGLWATEAKRDAKMSAREVAIVNTENIRQMDRSKRVTKNC 143
Db 308 QDLSSQTLRTAIIVAANAALKEQAQEOQLADCE-SFSDRRQDADBDVRIIQQDC 360

RESULT 4
US-08-895-590-2
; Sequence 2, Application US/08895590
; Patent No. 6207410
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian

APPLICANT: Zheng, Wei
APPLICANT: Dubald, Manuel Marcel Paul
TITLE OF INVENTION: Genes Encoding an Insect Calcium Channel
NUMBER OF SEQUENCES: 101
CORRESPONDENCE ADDRESS:
ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
STREET: 699 Prince Street
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/895,590
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/374,888
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NUMBER: 39,300
REFERENCE/DOCKET NUMBER: 022650-263
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-895-590-2

Query Match 10.5%; Score 80.5; DB 3; Length 2516;
Best Local Similarity 22.4%; Pred. No. 15;
Matches 39; Conservative 23; Mismatches 59; Indels 53; Gaps 8;
QY 3 CCGSRADAIEPRY--YESWTRETESTWLTYYDSALPSAATDSCPEAGLHAGVLEBGL 60
Db 207 CCGGGISAPPRRLTPEBAWLOQPQ-----NSVTSAGSTNSFSSG-----GR 249
QY 61 SSGVLRPAAPGCIANPEKKNK-----GTQC-----PNS 90
Db 250 DDNSGY--SAVGGDSSSSNSCNCIDITGDNSTLHGLGVGVCSFIADCDNSEDGDDPNN 307
QY 91 QNLSSGPL-TQKONGLWATEAKRDAKMSAREVAIVNTENIRQMDRSKRVTKNC 143
Db 308 QDLSSQTLRTAIIVAANAALKEQAQEOQLADCE-SFSDRRQDADBDVRIIQQDC 360

RESULT 5
US-09-539-879A-2
; Sequence 2, Application US/09539879A
; Patent No. 6436627
; GENERAL INFORMATION:
; APPLICANT: Hall, Linda M.
; APPLICANT: Ren, Dejian
; APPLICANT: Zheng, Wei
; APPLICANT: Dubald, Manuel Marcel Paul
; TITLE OF INVENTION: Genes Encoding an Invertebrate Alpha
; TITLE OF INVENTION: Calcium Channel Subunit
; NUMBER OF SEQUENCES: 57
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS, LLP
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA


```
ZIP: 22314-3187
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/539,879A
FILING DATE: 31-Mar-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/09/111,865
FILING DATE: <Unknown>
APPLICATION NUMBER: US 08/374,077
FILING DATE: 19-JAN-1995
ATTORNEY/AGENT INFORMATION:
NAME: McGowan, Malcolm M.
REGISTRATION NO: 39,300
REFERENCE/DOCKET NUMBER: 022650-264
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-836-6620
TELEFAX: 703-836-2021
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 2516 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 2:
US-09-539-879A-2

Query Match      10.5%; Score 80.5; DB 4; Length 2516;
Best Local Similarity 22.4%; Pred. No. 15;
Matches 39; Conservative 23; Mismatches 59; Indels 53; Gaps 8;

QY 3 CGGSRADAEPRY--YESWTRETESTWLTYYTDSALPSAAATDSCPEAGLHAGVLEDL 60
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 207 CGCGGSAPEPRRTPEANQLPQ-----NSVTSAGSTSSFSGG-----GR 249

QY 61 SSNGVLRPAAPGGINPEKKMNC-----GTQC-----PNS 90
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 250 DNNSSY--SAVGDSSSNSCNDITGDNSTLHGLGVGVCSFIADCDNSEDDEDPNN 307

QY 91 QNLSGPEL-TQKNGMLATEAKRDARMSAREVAIVTEIRQMDSKVTKNC 143
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 308 QDLSSQTLRTAIVAAVAAAKQAEQSLADCE-SFSDRQDADDEVRITQDC 360

RESULT 6
US-09-949-016-7117
Sequence 7117, Application US/09949016
Patent No. 6812339
GENERAL INFORMATION:
APPLICANT: VENTER, J. Craig et al.
TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
FILE REFERENCE: CL001307
CURRENT APPLICATION NUMBER: US/09/949,016
CURRENT FILING DATE: 2000-04-14
PRIOR APPLICATION NUMBER: 60/241,755
PRIOR FILING DATE: 2000-10-20
PRIOR APPLICATION NUMBER: 60/237,768
PRIOR FILING DATE: 2000-10-03
PRIOR APPLICATION NUMBER: 60/231,498
PRIOR FILING DATE: 2000-09-08
NUMBER OF SEQ ID NOS: 207012
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 7117
LENGTH: 694
TYPE: PRT
ORGANISM: Human
US-09-949-016-7117
```

```
Query Match      10.4%; Score 79.5; DB 4; Length 694;
Best Local Similarity 30.9%; Pred. No. 3.1;
Matches 29; Conservative 6; Mismatches 34; Indels 25; Gaps 3;

QY 31 TQSDALPSAAATDSCPEAGLH-----AGVLEDLSSNGVLRPAAP-- 71
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 233 TQSEVSGSPAKGSKVHNGHPQSPAVPPTYPSGPPPAASLSTRTGNGGVAPAAPG 292

QY 72 --GGIANPEKKMNCGTQCPSNQLS---SGPLT 99
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 293 ALGPKASPAASHNSGTPEYQAQVAVAPAPSGPST 326

RESULT 7
US-09-252-991A-18465
Sequence 18465, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 33142
SEQ ID NO 18465
LENGTH: 244
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-18465

Query Match      10.3%; Score 79; DB 4; Length 244;
Best Local Similarity 24.0%; Pred. No. 0.78;
Matches 42; Conservative 23; Mismatches 74; Indels 36; Gaps 6;

QY 2 CGGSRADAEPRYYESWTRETESTWLT-----YTDSALPSAAATDSCPEAGLHAGV 56
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 74 GQASARPARSPARSRSRSETPPARSSPSPGTHSSRLPARAGGTSSIGTASRLA 133

QY 57 EDGLSSNGVLRPAAPGGINPEKKMNC-----GTQCPSNQLS--- 92
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 134 GAVLRING--RPSASGIASPSAIPHCARPOAHQRPRRPAASGSSSRATRNDSQNP 191

QY 93 --LSSGPELTQKNGMLATEAKRDARMSAREVAIVTEIRQMDSKVTKNC 145
  ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 192 CRLAGSGTSKTSVASSGVSKPRWRKRRTSARLT--IRQART-GTSKPAIN 243
```

```
RESULT 8
US-09-121-964-1
Sequence 1, Application US/09121964
Patent No. 6124447
GENERAL INFORMATION:
APPLICANT: Natori, Shunji
TITLE OF INVENTION: NOVEL ENZYME CATALYZING DEPHOSPHORYLATION
FILE REFERENCE: 32290-144753
CURRENT APPLICATION NUMBER: US/09/121,964
CURRENT FILING DATE: 1998-07-24
NUMBER OF SEQ ID NOS: 9
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 724
TYPE: PRT
ORGANISM: Sarcophaga peregrina
US-09-121-964-1

Query Match      10.2%; Score 78; DB 3; Length 724;
Best Local Similarity 27.7%; Pred. No. 4.9;
Matches 31; Conservative 8; Mismatches 39; Indels 34; Gaps 4;
```

```

QY      25 STLTLYTSDSLPSAALTDSGP--EAGELHAGVLEDEG-----SSNG 65
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      2  STTEIVTQKQIHDLELFEKKKFPAGSCGKSLBEDITNRCBPLFKKDYTLIEDMSNV 61

QY      66 LRPAABGIANPE-----KKNCGTQCPNSQNLSSGPLTQKQNGL 105
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
Db      62 LSPRYPGRIPLPEYEHGHTSTTPNTNTNLGTQATNTNASPQSG--KNNGGL 110
           |||  |||  |||  |||  |||  |||  |||  |||  |||  |||

```

```

RESULT 9
US-09-252-991A-32740
: Sequence 32740 Application US/09252991A
: Patent No. 6551795
: GENERAL INFORMATION:
: APPLICANT: Marc J. Rubenfield et al.
: TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
: TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
: FILE REFERENCE: 107196.136
: CURRENT APPLICATION NUMBER: US/09/252,991A
: CURRENT FILING DATE: 1999-02-18
: PRIOR APPLICATION NUMBER: US 60/074,788
: PRIOR FILING DATE: 1998-02-18
: PRIOR APPLICATION NUMBER: US 60/094,190
: PRIOR FILING DATE: 1998-07-27
: NUMBER OF SEQ ID NOS: 33142
: SEQ ID NO 32740
: LENGTH: 715
: TYPE: PR1
: ORGANISM: Pseudomonas aeruginosa
: US-09-252-991A-32740

```

Query Match 10.0%; Score 77; DB 4; Length 715;
Best Local Similarity 17.9%; Pred. No. 6.3;
Matches 30; Conservative 25; Mismatches 43; Indels 70; Gaps 5

```

Qy 39 AAATDSGPBEGELHAAVLVDGL-----SSNGVL----- 66
Db 417 AAARKSAVQLQGHHADDAENGIIHAAVVGQADQGLAGGEGILMHGQYHPHVRPSSEPPFA 476
Qy 67 -----RPAPGSIANPEKKNGCFQPCNSQNLSSGPLYTKQNGLMATAKRD 113
Db 477 PQRSCSHKRGSEBPCPELOPLASRIAAST-----ASAPASASNSLITSSSKRS 528
Qy 114 A-----KMSAREVAIVNTEN-----IRQMSRKRYVTNGCI 144
Db 529 ACSQANSATRIAIATRASMTHTGRPPKPLSNCALRIISLPASRIASAVV 576

```

RESULT 10
 US-09-071-035-176
 Sequence 176, Application US/09071035
 Patent No. 6448043
 GENERAL INFORMATION:
 APPLICANT: Gil H. Choi
 TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
 NUMBER OF SEQUENCES: 496
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 FILING DATE: US/09/071,035
 CLASSIFICATION:

```

? PRIOR APPLICATION DATA:
? APPLICATION NUMBER:
? FILING DATE:
? ATTORNEY/AGENT INFORMATION:
? NAME: A. Anders Brookes
? REGISTRATION NUMBER: 36,373
? REFERENCE/DOCKET NUMBER: P3369P2
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: (301) 309-8504
? TELEFAX: (301) 309-8512
? INFORMATION FOR SEQ ID NO: 176:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 286 amino acids
? TYPE: amino acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: protein
? OS-09-071-035-176

```

Query Match	9.8%	Score 75.5;	DB 4;	Length 286;
Best Local Similarity	27.1%;	Pred. No. 2.5;		
Matches 48;	Conservative 16;	Mismatches 72;	Indels 41;	Gaps 11;

QY CGG-----SRADIERPRRYESV-----TETETETWLTITDSDALPSAAAT--DSGEEAG 49
Db 1 CGGCKSTENTIDRSRBAESTTVESTKASATSSSK--ATTSSDAKSSGTTTADSKATAS 59
QY 50 GHAAGVLEGLSSNGVLRPAAPGGIANGPEKKCN--CGTQCPV---SQUSSCPITQKONGL 105
Db 60 STKEA-----ANNSSAEKQSPAKKANPDQDQNVOLNANMPGQGPQAIILISQTNMF 113
QY 106 W--ATEAKDARMS-----AREVALINTE--NIRQND-----RSKRVTKNCIN 145
Db 114 LTAATTSQADDQNNFRVLVYAEKEALPVNDADVNLQTPISSEFKTYGSDAEAKNAVN 170

RESULT 11
 US-09-071-035-174
 Sequence 174, Application US/09071035
 Patent No. 6448043
 GENERAL INFORMATION:
 APPLICANT: Gil H. Choi
 TITLE OF INVENTION: Enterococcus faecalis Polynucleotides and Polypeptides
 NUMBER OF SEQUENCES: 496
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Human Genome Sciences, Inc.
 STREET: 9410 Key West Avenue
 CITY: Rockville
 STATE: Maryland
 COUNTRY: USA
 ZIP: 20850
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
 COMPUTER: HP Vectra 486/33
 OPERATING SYSTEM: MSDOS version 6.2
 SOFTWARE: ASCII Text
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/071,035
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER:
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: A. Anders Brookes
 REGISTRATION NUMBER: 36,373
 REFERENCE/DOCKET NUMBER: PB369P2
 TELECOMMUNICATION INFORMATION:
 TELEFAX: (301) 309-8512
 INFORMATION FOR SEQ. ID NO: 174:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 305 amino acids

TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-071-035-174

Query Match
Best Local Similarity 27.1%; Pred. No. 2.7;
Matches 48; Conservative 16; Mismatches 72; Indels 41; Gaps 11;

QY 3 CGG-----SRADALEPRYSW---TRETSTWLYTSDALPSAAT-DSGEAG 49
DB 20 CGGKSTENTDSRSSAESTTVESTKATKSSSK-ATTKSDAKSGTTTADSKATAS 78
QY 50 GLHAGVLEDLSNGVLRPAAPGIANPEKKM-CGTQCN---SQWLSGGLTKONGL 105
DB 79 STRKA-----ANNSSAEKQSPAKNAPDDQANQVNLQALNMPGQGLPPAIIITSGTNF 132
QY 106 W-ATEAKRDARMS-----AREVAIVTE-NIRQMD-----RSKRVTKNCIN 145
DB 133 LTAATTSQADQNNFRVLYAEKEAIPVNDARVQULTFISFEKKTYSDAEAKVAVN 189

RESULT 12

US-09-408-647A-2
Sequence 2, Application US/09408647A
Patent No. 6399858

GENERAL INFORMATION:
APPLICANT: Kobayashi, Donald
TITLE OF INVENTION: Chitinase Gene from Stenotrophomonas
FILE REFERENCE: Rut-Cook 98-0090
CURRENT APPLICATION NUMBER: US/09/408,647A
CURRENT FILING DATE: 1999-08-26
PRIOR APPLICATION NUMBER: 60/098,036
PRIOR FILING DATE: 1998-08-27
NUMBER OF SEQ ID NOS: 2
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 2
LENGTH: 700
TYPE: PRT
ORGANISM: Stenotrophomonas maltophilia
FEATURE:
NAME/KEY: SIGNAL
LOCATION: (1)...(41)
NAME/KEY: DOMAIN
LOCATION: (196)...(290)
NAME/KEY: DOMAIN
LOCATION: (330)...(483)
PUBLICATION INFORMATION:
DATABASE ACCESSION NUMBER: Genbank No. 6399858 AF014950
DATABASE ENTRY DATE: 1997-09-23
US-09-408-647A-2

Query Match
Best Local Similarity 22.5%; Pred. No. 9.1;
Matches 29; Conservative 15; Mismatches 58; Indels 27; Gaps 2;

QY 26 TWLYTSDALPSAATDSGPEAGLHAGVLEDLSNGVLRPAAPGIANPEKKM--- 81
DB 159 TWANASGSHTFRAVATDNNATVTSATVVTYASNDTTPPSVPGGLSPKXTATTN 218
QY 82 -----NCG-----TQCNQNLSSGGLTKONGLMATEAKRDARMS 118
DB 219 LVMSAATDNGSGGVAGDYVRNGSLVSGSPSAQTQYDGLTASTAYTYVRARDNAGNAS 278
QY 119 AREVAIVNT 127
DB 279 ASGGSISVT 287

RESULT 13
US-09-328-352-8115

Sequence 8115, Application US/09328352
Patent No. 6562958
GENERAL INFORMATION:
APPLICANT: Gary L. Breton et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
FILE REFERENCE: GTC99-03PA
CURRENT APPLICATION NUMBER: US/09/328,352
CURRENT FILING DATE: 1999-06-04
NUMBER OF SEQ ID NOS: 8252
SEQ ID NO 8115
LENGTH: 726
TYPE: PRT
ORGANISM: Acinetobacter baumannii
US-09-328-352-8115

Query Match
Best Local Similarity 33.8%; Pred. No. 14;
Matches 26; Conservative 8; Mismatches 23; Indels 20; Gaps 4;

QY 2 GCGGSRADALEPRYSWTRTESTWLYTSDALPSA--AATDSC-----PEAGLHAG 54
DB 181 GFGAGREDVWEPNDVNWGDKE--WLAHNSBALAGSNLAATEMGLIYVNP----- 231
QY 55 VLEDGLSSNGVLRPAAP 71
DB 232 ---GPQASGDPRSAP 244

RESULT 14

US-09-618-425-9
Sequence 9, Application US/09618425
Patent No. 6475744

GENERAL INFORMATION:
APPLICANT: Reppert, Steven M.
APPLICANT: Weaver, David R.
APPLICANT: Zylka, Mark
APPLICANT: Jin, Xiaowei
APPLICANT: Kume, Kazuhiko
APPLICANT: Sitram, Sathyanarayanan
APPLICANT: Shearman, Lauren
TITLE OF INVENTION: METHODS FOR IDENTIFYING COMPOUNDS WHICH
MODULATE CIRCADIAN RHYTHM
FILE REFERENCE: 00786-428001
CURRENT APPLICATION NUMBER: US/09/618,425
CURRENT FILING DATE: 2000-07-18
PRIOR APPLICATION NUMBER: 60/203,005
PRIOR FILING DATE: 2000-05-10
PRIOR APPLICATION NUMBER: 60/145,363
PRIOR FILING DATE: 1999-07-22
NUMBER OF SEQ ID NOS: 13
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 9
LENGTH: 1113
TYPE: PRT
ORGANISM: Mus musculus
US-09-618-425-9

Query Match
Best Local Similarity 25.0%; Pred. No. 27;
Matches 38; Conservative 22; Mismatches 70; Indels 22; Gaps 7;

QY 1 MCGGGS-----RADALEPRYSWTRTESTWLYTSDALPSAATDSGPEAGL 51
DB 860 LGAAGSSELAPLVPAAPNPDEPTTSGHSORRYEENNEAH--SHELFPISRSSP-----L 913
QY 52 HAGVLEDGLSSNGVLRPAAPGIANPEKKMNC--GTQCNQNLSSGGL---TQKONGLM 106
DB 914 QNLTLQEMPAPESADAVRRG-AGPDARHCVTGTGSGSRSRRTSGELATATAHQESAA 972
QY 107 ATEAKRDARMSAREVAIVNTEN-IRQMDRSK 137
DB 973 ASGGSASITVPSSTDVASEVSENRQRPPDRQR 1004

RESULT 15

US-09-538-092-723
 ; Sequence 723, Application US/09538092
 ; Patent No. 6753314
 ; GENERAL INFORMATION:
 ; APPLICANT: Glot, Loic
 ; APPLICANT: Mansfield, Traci A.
 ; TITLE OF INVENTION: Protein-Protein Complexes and Method of Using Same
 ; FILE REFERENCE: 15966-542
 ; CURRENT APPLICATION NUMBER: US/09/538,092
 ; CURRENT FILING DATE: 2000-03-29
 ; PRIOR APPLICATION NUMBER: 60/127,352
 ; PRIOR FILING DATE: 1999-04-01
 ; PRIOR APPLICATION NUMBER: 60/178,965
 ; PRIOR FILING DATE: 2000-02-01
 ; NUMBER OF SEQ ID NOS: 1387
 ; SOFTWARE: CuraPatSeqFormatter Version 0.9
 ; SEQ ID NO 723
 ; LENGTH: 346
 ; TYPE: PRT
 ; ORGANISM: Saccharomyces cerevisiae
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (0)...(0)
 ; OTHER INFORMATION: Polypeptide Accession Number YOR007C
 US-09-538-092-723

Query Match 9.6%; Score 73.5; DB 4; Length 346;
 Best Local Similarity 29.4%; Pred. No. 5.6;
 Matches 37; Conservative 11; Mismatches 51; Indels 27; Gaps 6;

QY 5 GSRADAIEPRYVESWTRTE-STWLTFT-----DSDALPSAATDSG-PEAGGLHAGVL 56
 DB 201 GDNATEAMKRDYIESAKKVEQSINLEKTVPEQSRDADVDASQGSAGSLPDLGSLGGGL 260
 QY 57 EDGLSSNGVLRPA-----PGGIANPEKKM-----CGTCCPNSQNTLSSGPL 98
 DB 261 -GGLMNNPQLMQAAGQMGMSNPGAMQNIQGMMDPSIRQMAEGFASGGTPTNLSDLNNPDA 319
 QY 99 TOKONG 104
 DB 320 LRNMAG 325

Search completed: April 12, 2005, 15:43:30
 Job time : 27.6667 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:37:30 ; Search time 75.6667 Seconds
(without alignments)
636.206 Million cell updates/sec

Title: US-10-705-716a-8

Perfect score: 767

Sequence: 1 MCGGSRADAIERYESWT.....VTENIRQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1418010 seqs, 33197259 residues

Total number of hits satisfying chosen parameters: 1418010

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :

Published Applications AA:*

1: /cgn2_6/prodata/1/pubppa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/1/pubppa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/1/pubppa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/1/pubppa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/1/pubppa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/1/pubppa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/1/pubppa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/1/pubppa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/1/pubppa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/1/pubppa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/1/pubppa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/1/pubppa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
14: /cgn2_6/prodata/1/pubppa/US10_PUBCOMB.pep:*
15: /cgn2_6/prodata/1/pubppa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
17: /cgn2_6/prodata/1/pubppa/US10_NEW_PUB.pep:*
18: /cgn2_6/prodata/1/pubppa/US11_NEW_PUB.pep:*
19: /cgn2_6/prodata/1/pubppa/US60_NEW_PUB.pep:*
20: /cgn2_6/prodata/1/pubppa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	649	84.6	145	US-10-293-239-17	Sequence 17, Appl
2	649	84.6	145	US-10-177-390-30	Sequence 30, Appl
3	621.5	81.0	180	US-10-293-239-19	Sequence 19, Appl
4	472.5	61.6	149	US-10-293-239-21	Sequence 21, Appl
5	288	37.5	73	US-10-293-239-20	Sequence 20, Appl
6	278	36.2	54	US-10-293-239-18	Sequence 18, Appl
7	272	35.5	80	US-10-293-239-22	Sequence 22, Appl
8	140	18.3	25	US-10-293-239-37	Sequence 37, Appl
9	99	12.9	18	US-10-293-239-35	Sequence 35, Appl
10	86	11.2	1597	US-10-437-963-180225	Sequence 180225,
11	84	11.0	147	US-10-767-701-48905	Sequence 48905, A
12	84	11.0	1433	US-09-801-368-60	Sequence 60, Appl
13	84	11.0	1433	US-10-149-310-112	Sequence 112, App

14	82.5	10.8	274	US-10-425-114-65588	Sequence 65588, A
15	82.5	10.8	609	US-10-437-963-144983	Sequence 144983,
16	82	10.7	455	US-10-424-599-193656	Sequence 193656,
17	82	10.7	1479	US-10-437-963-180217	Sequence 180217,
18	81	10.6	219	US-10-156-761-13447	Sequence 13447, A
19	80.5	10.5	600	US-10-408-765A-1945	Sequence 1945, Ap
20	80	10.4	528	US-10-437-963-173618	Sequence 173618,
21	80	10.4	1108	US-09-949-029-108	Sequence 108, App
22	79.5	10.4	996	US-10-437-963-114095	Sequence 114095,
23	79.5	10.3	876	US-09-801-368-44	Sequence 44, Appl
24	78.5	10.2	1274	US-10-437-963-180220	Sequence 180220,
25	78	10.2	1491	US-10-437-963-140197	Sequence 140197,
26	77.5	10.1	670	US-10-156-761-14107	Sequence 14107, A
27	77	10.0	268	US-10-369-493-8340	Sequence 8340, Ap
28	76.5	10.0	320	US-10-425-114-43318	Sequence 43318, A
29	76.5	10.0	426	US-10-369-493-3326	Sequence 3326, Ap
30	76	9.9	391	US-10-437-963-177907	Sequence 177907,
31	75.5	9.8	276	US-10-437-963-158275	Sequence 158275,
32	75.5	9.8	286	US-09-071-035-176	Sequence 176, App
33	75.5	9.8	286	US-10-206-576-176	Sequence 176, App
34	75.5	9.8	286	US-10-912-362-176	Sequence 176, App
35	75.5	9.8	305	US-09-071-035-174	Sequence 174, App
36	75.5	9.8	305	US-10-206-576-174	Sequence 174, App
37	75.5	9.8	305	US-10-912-362-174	Sequence 174, App
38	75.5	9.8	373	US-10-437-963-104701	Sequence 104701,
39	75	9.8	167	US-10-437-963-146805	Sequence 46805, A
40	75	9.8	604	US-10-437-963-163101	Sequence 163101,
41	74.5	9.7	296	US-10-437-963-193696	Sequence 193696,
42	74.5	9.7	580	US-10-282-122A-73164	Sequence 73164, A
43	74.5	9.7	601	US-10-282-122A-74859	Sequence 74859, A
44	74.5	9.7	601	US-10-282-122A-76267	Sequence 76267, A
45	74	9.6	332	US-10-437-963-168390	Sequence 168390,

ALIGNMENTS

RESULT 1					
US-10-293-239-17					
Sequence 17, Application US/10293239					
Publication No. US20030119043A1					
GENERAL INFORMATION:					
APPLICANT: Tanner, Stephan					
TITLE OF INVENTION: BAAC expression as a diagnostic marker for acute leukemia					
FILE REFERENCE: 22727/04101					
CURRENT APPLICATION NUMBER: US/10/293,239					
CURRENT FILING DATE: 2002-11-12					
PRIOR APPLICATION NUMBER: US 60/348,210					
PRIOR FILING DATE: 2001-11-09					
NUMBER OF SEQ ID NOS: 39					
SOFTWARE: PatentIn version 3.1					
SEQ ID NO 17					
LENGTH: 145					
TYPE: PRT					
ORGANISM: Homo sapiens					
US-10-293-239-17					
Query Match					
Best Local Similarity 84.6%; Score 649; DB 14; Length 145;					
Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;					
QY					
1 MCGGSRADAIERYESWTRETSTWLTYSDDALPSAATDSGPGAGLHGVLEDC 60					
DB					
1 MCGGSRADAIERYESWTRETSTWLTYSDDALPSAATDSGPGAGLHGVLEDC 60					
QY					
61 SSNGVLRPAAPGGIAPPEKMKNGCTCPNSQNSASPELTKQKGLWATEKRPKXMSAR 120					
DB					
61 PSNGVLRPAAPGGIAPPEKMKNGCTCPNSQNSASPELTKQKGLWATEKRPKXMSAR 120					
QY					
121 EVAINVTENIRQMDRSKRVTKNCIN 145					
DB					
121 EVAINVTENIRQMDRSKRVTKNCIN 145					

```

RESULT 2
US-10-177-390-30
; Sequence 30, Application US/10177390
; Publication No. US20030143743A1
; GENERAL INFORMATION:
; APPLICANT: Schuler, Gerold
; APPLICANT: N.V. Antwerp Innovatiecentrum
; TITLE OF INVENTION: Improved Transfection of Eucaryotic Cells with Linear
; TITLE OF INVENTION: Polynucleotides by Electroporation
; FILE REFERENCE: 021503wo/3H/ml
; CURRENT APPLICATION NUMBER: US/10/177,390
; CURRENT FILING DATE: 2002-06-20
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 30
; LENGTH: 145
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-177-390-30

```

Query Match	84.6%	Score 649	DB 14	Length 145
Best Local Similarity	83.4%	Pred. No. 8.6e-60		
Matches 121	Conservative 10	Mismatches 14	Indels 0	Gaps 0
Qy	1	MGCGGSRDAIEPRYYESWTRTESTWLTYYTSDALPSAAATDPSGEAGGLHAGVLEDGL	60	
Db	1	MGCGGSRDAIEPRYYESWTRTESTWLTYYTSDAPPSAAAPDSGEAGGLHSGMLEDGL	60	
Qy	61	SSNGYLPAAPAGGIANPEKKMGCTGCPNSQNLSSGPTLQKONGLWATEAKDRAKMSAR	120	
Db	61	PSNGVPRSTAPEGIINPEKKTNCETGCPHPSGLSSGPLLQKONGLTTEAKDRAKMPAK	120	
Qy	121	EVAIVTENIKOMDRSKRYTKCIN	145	
Db	121	EVTIVTDSIQOMDRSRRTTKCVN	145	

```

RESULT 3
US-10-293-239-19
: Sequence 19, Application US/10293239
: Publication No. US20030119043A1
: GENERAL INFORMATION:
: APPLICANT: Tanner, Stephan
: APPLICANT: de la Chapelle, Albert
: TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
: FILE REFERENCE: 22727/04101
: CURRENT APPLICATION NUMBER: US/10/293,239
: PRIOR FILING DATE: 2002-11-12
: PRIOR APPLICATION NUMBER: US 60/348,210
: PRIOR FILING DATE: 2001-11-09
: NUMBER OF SEQ ID NOS: 39
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 19
: LENGTH: 180
: TYPE: PRT
: ORGANISM: Homo sapiens
US-10-293-239-19

```

	Query Match	81.0%;	Score 621.5; DB 14; Length 180;
	Best Local Similarity	67.2%;	Pred. No. 8-6e-57;
	Matches 121; Conservative	10;	Mismatches 14; Indels 35; Gaps 1
QY	1 MGCGSRADAIPEPRYESWTRTESTTWLTYTSDALPSAAATDSGEGAGLA-----	53	
Dy	1 MGCGSRADALEPRTESWTRETSWLTYTSDAPPSAAPPDGGEGAGLSVLEAEKS	60	
OY	54 -----GVLDGLSSNGVLRLPAAGJANPEKKNNCGT	85	
Dd	61 KIRAPTDSVBDEGLFSASIKMAPLAVFSGHMLDGCLPSNVGPSTAGGIINPEKKTNCET	120	
OY	86 QCPNSONLISSGPLTKOKNGLMATEAKRDAKRMSAREVALINVENTNRQMORSKRVTVNCCTN	145	

[illegible]

```

RESULT 4
US-10-293-239-21
; Sequence 21, Application US/10293239
; Publication No. US20030119043b1
; GENERAL INFORMATION:
; APPLICANT: Tanneir, Stephan
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukaemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293.239
; CURRENT FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; PRIOR FILING DATE: 2001-11-09
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 21
; LENGTH: 149
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-21

```

	Query Match	61.6%	Score 472.5;	DB 14;	Length 149;				
	Best Local Similarity	65.3%;	Pred. No. 2.7e-41;						
	Matches	94;	Conservative	3; Mismatches 12;	Indels 35; Gaps 1				
Qy	1	MGGCGSRADAI	EPRYYESWTR	ETESTTWLT	YDSDALPSAAATD	GCPAGGLHA-----	53		
Dd	1	MGGGGRDAI	EPRYYESWTR	ETESTTWLT	YDSDALP	SAAAAPDSGPA	AGLHSLVLEAKS	60	
Qy	54	-----	GVLEDGL	SSNGVL	RPAAPGI	ANPEKON	CGT	85	
Dd	61	KIKAPTDSV	DEGLPSASMA	PLAVFHS	HGMLEDGL	PSNGVPR	STAPGCI	INPEKTCVCT	120
Qy	86	QCPSNQUS	SGPVTOKONG	LIMATE	109				
Dd	121	QCNPQSLS	SGPVTOKONG	LQTTE	144				

```

RESULT 5
US-10-293-239-20
Sequence 20, Application US/10293239
Publication No. US20030119043A1
GENERAL INFORMATION:
APPLICANT: Tanneer, Stephan
APPLICANT: de la Chapelle, Albert
TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
FILE REFERENCE: 23727/04101
CURRENT APPLICATION NUMBER: US/10/293,239
CURRENT FILING DATE: 2002-11-12
PRIOR APPLICATION NUMBER: US 60/348,210
PRIOR FILING DATE: 2001-11-09
NUMBER OF SEQ ID NOS: 39
SOFTWARE: PatentIn version 3.1
SEQ ID NO 20
LENGTH: 73
TYPE: PRT
ORGANISM: Homo sapiens
US-10-293-239-20

```

Query Match	Similarity	Score	DB	Length
Beet	Local	91.4%	Pred. No. 2.2e22	
Matches	53	Conservative	2	Indels 0; Gaps 0

Qy	1	MGCGGSRADAI	BPYYE	SWT	RETES	TW	LT	YD	SAL	PS	AA	T	D	S	G	P	A	G	L	H	A	G	L	A	G	L	E	D	58
bb	1	MGCGGSRADAI	BPYYE <td>SWT</td> <td>RETES</td> <td>TW <td>LT <td>YD <td>SAL</td> <td>PS <td>AA</td> <td>T</td> <td>D</td> <td>S</td> <td>G <td>P <td>A</td> <td>G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td></td></td></td></td></td></td></td>	SWT	RETES	TW <td>LT <td>YD <td>SAL</td> <td>PS <td>AA</td> <td>T</td> <td>D</td> <td>S</td> <td>G <td>P <td>A</td> <td>G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td></td></td></td></td></td></td>	LT <td>YD <td>SAL</td> <td>PS <td>AA</td> <td>T</td> <td>D</td> <td>S</td> <td>G <td>P <td>A</td> <td>G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td></td></td></td></td></td>	YD <td>SAL</td> <td>PS <td>AA</td> <td>T</td> <td>D</td> <td>S</td> <td>G <td>P <td>A</td> <td>G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td></td></td></td></td>	SAL	PS <td>AA</td> <td>T</td> <td>D</td> <td>S</td> <td>G <td>P <td>A</td> <td>G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td></td></td></td>	AA	T	D	S	G <td>P <td>A</td> <td>G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td></td></td>	P <td>A</td> <td>G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td></td>	A	G <td>L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td></td>	L <td>H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td></td>	H <td>A</td> <td>G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td></td>	A	G <td>L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td></td>	L <td>A</td> <td>G <td>L <td>E <td>D</td> <td>58</td> </td></td></td>	A	G <td>L <td>E <td>D</td> <td>58</td> </td></td>	L <td>E <td>D</td> <td>58</td> </td>	E <td>D</td> <td>58</td>	D	58

RESULT 6

```
US-10-293-239-18
; Sequence 18, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 18
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-18

Query Match      36.2%; Score 278; DB 14; Length 54;
Best Local Similarity 94.4%; Pred. No. 1.7e-21;
Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 MCGGSRADAIEPRYESWTRETESTWLTITDSDALPSAATDGPAGGLHAG 54
Db      1 MCGGSRADAIEPRYESWTRETESTWLTITDSDAPPSAAPPDGPAGGLHSG 54

RESULT 7
US-10-293-239-22
; Sequence 22, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 22
; LENGTH: 80
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-22

Query Match      35.5%; Score 272; DB 14; Length 80;
Best Local Similarity 94.3%; Pred. No. 1.2e-20;
Matches 50; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY      1 MCGGSRADAIEPRYESWTRETESTWLTITDSDALPSAATDGPAGGLHAG 53
Db      1 MCGGSRADAIEPRYESWTRETESTWLTITDSDAPPSAAPPDGPAGGLHGS 53

RESULT 8
US-10-293-239-37
; Sequence 37, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 37
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-37

Query Match      18.3%; Score 140; DB 14; Length 25;
Best Local Similarity 100.0%; Pred. No. 1.8e-07;
Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      7 RADAIERPRYESWTRETESTWLTIT 31
Db      1 RADAIERPRYESWTRETESTWLTIT 25

RESULT 9
US-10-293-239-35
; Sequence 35, Application US/10293239
; Publication No. US20030119043A1
; GENERAL INFORMATION:
; APPLICANT: Tanner, Stephan
; APPLICANT: de la Chapelle, Albert
; TITLE OF INVENTION: BALC expression as a diagnostic marker for acute leukemia
; FILE REFERENCE: 22727/04101
; CURRENT APPLICATION NUMBER: US/10/293,239
; PRIOR FILING DATE: 2002-11-12
; PRIOR APPLICATION NUMBER: US 60/348,210
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
; LENGTH: 18
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-239-35

Query Match      12.9%; Score 99; DB 14; Length 18;
Best Local Similarity 100.0%; Pred. No. 0.0023;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      9 DAIEPRYESWTRETEST 26
Db      1 DAIEPRYESWTRETEST 18

RESULT 10
US-10-437-963-180225
; Sequence 180225, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukhariou, Andrey A.
; APPLICANT: Bardazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; PRIOR FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 180225
; LENGTH: 1597
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_77610C.1.pcp
US-10-437-963-180225
```

RESULT 12
US-09-801-368-60
Sequence 60, Application US/09801368
Patent No. US20020128250A1
GENERAL INFORMATION:
APPLICANT: Busby, Robert
APPLICANT: Cali, Brian
APPLICANT: Hecht, Peter
APPLICANT: Holtzman, Doug
APPLICANT: Madden, Kevin
APPLICANT: Maxon, Mary
APPLICANT: Milne, Todd
APPLICANT: No. US20020128250A1man, Thea
APPLICANT: Royer, John
APPLICANT: Salama, Sofie
APPLICANT: Sherman, Amir
APPLICANT: Silva, Jeff
APPLICANT: Summers, Eric
TITLE OR INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
FILE REFERENCE: 109272.147
CURRENT APPLICATION NUMBER: US/09/801,368
CURRENT FILING DATE: 2001-03-07
PRIOR APPLICATION NUMBER: US 09/487,558
PRIOR FILING DATE: 2000-01-19
PRIOR APPLICATION NUMBER: US 60/160,587

```

RESULT 14
US-10-425-114-65588
; Sequence 65588, Application US/10425114
; Publication NO. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jiongong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with

```


Search completed: April 12, 2005, 16:03:43
Job time : 76.6667 secs

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(5313)B
/ CURRENT APPLICATION NUMBER: US/10/425,114
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 73128
/ SEQ ID NO 65588
/ LENGTH: 274
/ TYPE: PRT
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: LIB4863-011-Fl12_Fl1.pep
US-10-425-114-65588

Query Match 10.8%; Score 82.5; DB 15; Length 274;
Best Local Similarity 34.4%; Pred. No. 4;
Matches 31; Conservative 8; Mismatches 32; Indels 19; Gaps 4;

QY 32 DSDALPSAATDGGPAGGLHAGVLDGLSNGVLRPAAPGFI--ANPEKKMNCCTGCP 88
DB 147 DSDDAFAAAAHDDQPAVAVVAAAGLGCGNSSSG-LPPAAGATAFAEPISLSLGLPLP 205
QY 89 -----NSQNLSSGFLTQKONG 104
DB 206 AAPAPAAADDESRNSQGQAS-PLLEGGEG 234

RESULT 15
US-10-437-963-144983
/ Sequence 144983, Application US/10437963
/ Publication No. US2004012343A1
/ GENERAL INFORMATION:
/ APPLICANT: La Rosa, Thomas J.
/ APPLICANT: Kovacic, David K.
/ APPLICANT: Zhou, Yihua
/ APPLICANT: Cao, Yongwei
/ APPLICANT: Wu, Wei
/ APPLICANT: Boukharov, Andrey A.
/ APPLICANT: Barbazuk, Brad
/ APPLICANT: Li, Ping
/ TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
/ FILE REFERENCE: 38-21(53221)B
/ CURRENT APPLICATION NUMBER: US/10/437,963
/ CURRENT FILING DATE: 2003-05-14
/ NUMBER OF SEQ ID NOS: 204966
/ SEQ ID NO 144983
/ LENGTH: 609
/ TYPE: PRT
/ ORGANISM: Oryza sativa
/ FEATURE:
/ NAME/KEY: unsure
/ LOCATION: (1)..(609)
/ OTHER INFORMATION: unsure at all Xaa locations
/ FEATURE:
/ OTHER INFORMATION: Clone ID: PAT_MRT4530_45746C.1.pap
US-10-437-963-144983

Query Match 10.8%; Score 82.5; DB 16; Length 609;
Best Local Similarity 25.7%; Pred. No. 11;
Matches 38; Conservative 21; Mismatches 68; Indels 21; Gaps 5;

QY 8 ADALIERRYIESWTR-----ETESTWLTYYTSDALPSAATDGGPAGGL 51
DB 20 ADDEEEEBEATVRPVEVASKDQEEVEEBEEDSDVDALDMLDAE-GPDGSGR 78
QY 52 HAGVLDGLSNGVLRPAAPGFI-ANPEKKMNCCTGCPNSQNLSSGFLTQKON--GLMAT 108
DB 79 PAAAPTAAGGAARPNHAGVLSRPFQPISNRTQ-KLASHIRATPLEWEGRMVAGMS 137
QY 109 EAKRDAKMSAREVAIVNTENIRQMDRS 136
DB 138 NSVTAIKRDISIRETAIKTRNTEKADRA 165

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:28:59 ; Search time 21.6667 Seconds
(without alignments)
643.912 Million cell updates/sec

Title: US-10-705-716a-8

Sequence: 1 MCGGSGRADIEPRRYESWT.....VTENINQMDRSKRVTKNCIN 145

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87.5	11.4	778	2 A35970	erythrocyte-binding
2	84	11.0	365	1 GNVSSC	genome polyprotein
3	84	11.0	1433	2 S54587	CAT8 protein - yea
4	79.5	10.4	571	2 T43456	hypothetical prote
5	79.5	10.4	1063	2 T03743	bifocal protein -
6	79.7	10.3	876	1 A57988	regulatory protein
7	78.5	10.2	600	2 C69899	conserved hypochet
8	78	10.2	269	2 T37073	hypothetical prote
9	77.5	10.1	848	2 G90128	elongation factor
10	77.5	10.1	2559	2 T09144	probable guanine n
11	77	10.0	781	2 T41551	hypothetical prote
12	76	9.9	516	2 JB0301	inulinase (EC 3.2.
13	76	9.9	645	2 S19156	serotonin receptor
14	74.5	9.7	601	2 AH0784	probable transpor
15	74	9.6	1113	2 T14260	period protein per
16	73.5	9.6	281	2 T02813	hypothetical prote
17	73.5	9.6	346	2 S61991	hypothetical prote
18	73.5	9.6	435	2 AG1028	preillin [imported
19	72.5	9.5	451	2 S13337	tubulin alpha-2 ch
20	72.5	9.5	851	2 H69775	hypothetical prote
21	72.5	9.5	1050	2 S54640	KCS1 protein - yea
22	72	9.4	424	2 T01383	GMPase-activating
23	72	9.4	4957	2 T03455	AKR protein - huma
24	72	9.4	5262	2 T03454	AKR protein - huma
25	71.5	9.3	3488	2 T34418	hypothetical prote
26	71	9.3	652	2 G82962	hypothetical prote
27	71	9.3	1840	2 T30250	GPI protein - mous
28	71	9.3	5327	2 T13564	microtubule-associ
29	70.5	9.2	251	2 C75521	cytochrome-related

30	70.5	9.2	351	2 A56387	helix-loop-helix t
31	70.5	9.2	381	2 A11846	aulfate-binding pr
32	70.5	9.2	502	2 S28080	gag polyprotein
33	70.5	9.2	512	2 E95902	probable sugar ABC
34	70.5	9.2	1343	2 AF0611	cell division prot
35	70.5	9.2	1996	2 F71405	probable TWV resis
36	70.5	9.2	3623	2 T08618	intrinsic factor-B
37	70	9.1	368	1 TVMSML	transforming prote
38	70	9.1	419	2 A36109	pregnancy-specific
39	70	9.1	475	2 C70966	hypothetical prote
40	70	9.1	500	2 P90133	t-complex protein
41	70	9.1	614	2 T06741	hypothetical prote
42	70	9.1	1573	2 T50113	3-dehydroquinase s
43	69.5	9.1	284	2 I51172	transcription fac
44	69.5	9.1	346	2 AF3161	phage-related inte
45	69.5	9.1	367	1 MHCH	Ig mu chain C regi

ALIGNMENTS

RESULT 1

A35970
erythrocyte-binding protein - Plasmodium knowlesi
C/Species: Plasmodium knowlesi
C/Date: 14-Dec-1990 #sequence_revision 14-Dec-1990 #text_change 09-Jul-2004
C/Accession: A35970
R/Adams, J.H.; Hudson, D.E.; Torii, M.; Ward, G.E.; Wellms, T.E.; Aikawa, M.; Miller, J.
Cell 63, 141-153, 1990
A/Title: The Duffy receptor family of plasmodium knowlesi is located within the microne-
C/Reference number: A35970, PMID:91004213, PMID:2170017
A/Accession: A35970
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1-778 <ADA>
A/Cross-references: UNIPROT:P22545; GB:M68518; GB:M37513; NID:g160273; PID:g160274
C/Keywords: transmembrane protein

Query Match 11.4%; Score 87.5; DB 2; Length 778;
Best Local Similarity 31.8%; Pred. No. 3.1;
Matches 28; Conservative 13; Mismatches 44; Indels 3; Gaps 3;

QY 31 TDSDALPSAATSGSPGAGLHAGVLEDDGSSNGV-LRPAAPGGINPEKKNK-GTQCP 88
DB 352 TVSSDVPVSGKQSGSTSHLAGENGVRHNGTDPKEDKADPQKQDKQDT 411
QY 89 NSQNLSS-GPLTQKONGLMATEAKRDAK 115
DB 412 DRSQGLSPHTDERATLGETHMEKDE 439

RESULT 2

GNVSSC
genome polyprotein - sugarcane mosaic virus (strain SC) (fragment)
N/Contents: carboxyl end of nuclear inclusion protein b; coat protein
C/Species: sugarcane mosaic virus; SCMV
C/Date: 30-Jun-1993 #sequence_revision 30-Jun-1993 #text_change 09-Jul-2004
R/Frenkel, M.J.; Jilka, J.M.; McKern, N.M.; Strike, P.M.; Clark Jr., J.M.; Shukla, D.D.
J. Gen. Virol. 72, 237-242, 1991
A/Title: Unexpected sequence diversity in the amino-terminal ends of the coat proteins
A/Reference number: PH0207, PMID:91132116; PMID:1993866
A/Accession: PH0207
A/Molecule type: genomic RNA
A/Residues: 1-365 <FRB>
A/Cross-references: UNIPROT:P25242; GB:D00948; NID:g2222123; PIDW:BAA00796.1; PID:g222212
C/Superfamily: tobacco etch virus genome polyprotein
C/Keywords: coat protein; inclusion protein b (fragment) #status predicted <IPB>
F/1-52/Product: nuclear inclusion protein b (fragment) #status predicted <IPB>
F/53-365/Product: coat protein #status predicted <COA>

Query Match 11.0%; Score 84; DB 1; Length 365;
Best Local Similarity 21.6%; Pred. No. 2.7;

Matches	30, Conservative	23, Mismatches	80, Indels	6, Gaps	4,
QY	5	GSRADAIIPRYEASTRETETWLTYYTSDALPSAAATDSCBEAGLHAGVL--EDGLSS	62		
Db	21	GIKEEIEI-KYFKQAKDLPGYLEDYND-EVFHNGTVDAGAGGGGAGNATQPPATGAA	78		
QY	63	NGVLRLPAPPGGIANPERKMANCGTCQPNQNLSSGFLTYKQNGLWATEKRAKMSAAEV	122		
Db	79	QGGAGPPTTGAAGPPTTQ--GSQLPQGGATGGGGAGTGAAGTGTSVTGGQDKVDAGTT	136		
QY	123	AINVTENIRQMDRSKRYTK	141		
Db	137	GKITVYFKLAKMSKKRRLEPK	155		

RESULT 3

CAT8 protein yeast (*Saccharomyces cerevisiae*)
N:Alternate names: M2P8 protein; protein YMR802L.06c; protein YMR280C
C:Species: *Saccharomyces cerevisiae*
C:Date: 08-Jul-1995 #sequence revision 01-Sep-1995 #text_change 16-Aug-2004
C:Accession: S54587; S48234; S61595; S49498
R:Pearson, D.; Bowman, S.
submitted to the EMBL Data Library, May 1995
A:Reference number: S54582
A:Accession: S54587
A:Molecule type: DNA
A:Residues: 1-1433 <PEA>
A:Cross-references: UNIPROT:P39113; EMBL:Z49704; NID:g825540; PIDN:CAA69778.1; PID:g825540
A:Experimental source: strain AB972
R:Grzeslitza, D.
submitted to the EMBL Data Library, March 1994
A:Reference number: S48234
A:Accession: S48234
A:Molecule type: DNA
A:Residues: 1-746,'A',748-1433 <GRZ>
A:Cross-references: EMBL:X78344; NID:g559523; PIDN:CAA55139.1; PID:g559524
R:Boles, E.; Heltmann, C.; Zimmermann, F.K.
submitted to the EMBL Data Library, December 1995
A:Reference number: S61594
A:Accession: S61595
A:Molecule type: DNA
A:Residues: 1-302,'A',304-767,'A',769-998,1003-1007,'S',1009-1015,'O',1017-1018,'P',1020-1021
A:Cross-references: EMBL:X94215; NID:g1122900; PIDN:CAA63906.1; PID:e214033; PID:g1122900
A:Experimental source: strain ENY.WA-1A
C:Genetics:
A:Gene: SGD:CAT8; M2P8
A:Cross-references: SGD:S0004893; MIPS:YMR280C
A:Map position: 13R
C:Superfamily: GAL4 zinc binuclear cluster homology
C:Keywords: transmembrane protein
F:65-102/Domain: GAL4 zinc binuclear cluster homology <GAL4>
F:456-472/Domain: transmembrane #status predicted <TM1>
F:738-754/Domain: transmembrane #status predicted <TM2>

Query Match

Matches	32;	Conservative	15;	Mismatches	50;	Indels	12;	Gaps	4;
QY	36	LP\$AAATDSCGEAGLHAGVLEDEGL\$NSGVLRPAAPG\$INPEKKNNCGTCGCP\$SLOL\$S	95						
Db	927	LP\$ATTT\$SLKLP\$SOSK\$SL\$ENRQRT\$P\$NVK\$-----ENPE\$E\$YL\$G\$D\$NNNN\$N\$S\$E\$A	979						
QY	96	G--PLTQK\$NG\$LMATEAK\$RKAR\$MSARE\$VA\$IN\$T\$EN\$T\$--Q\$M\$DS\$K\$V\$T	140						
	980	GHS\$MT\$YTT\$TG\$N\$K\$R\$K\$Y\$E\$K\$D\$A\$K\$--N\$A\$K\$D\$G\$G\$S\$K\$E\$N\$H\$N\$F\$E\$N\$D\$T\$P\$K\$N\$S	1027						

RESULT 4

hypothetical protein DKFZp434L061.1 - human
C:Species: Homo sapiens (man)
C:Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004

```

C:Accession: 143456
R:Pousterka, A., Klein, M., Mewes, H.W., Gassenhuber, J., Wiemann, S.
submitted to the Protein Sequence Database, December 1999
A:Reference number: Z22516
A:Accession: T43456
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-571 <AAA>
A:Cross-references: UNIPROT:075175, EMBL:AL133647
A:Experimental source: adult testis, clone DKFp434L061
C:Genetics:
A:Note: DKFp434L061.1
C:Superfamily: hydroxyproline-rich glycoprotein

Query Match      10.4%; Score 79.5; DB 2; Length 571;
Best Local Similarity 30.9%; Pred. No. 12;
Matches 29; Conservative 6; Mismatches 34; Indels 25; Gaps 3

QY      31 TDSDALPAAALDSCGEAAGLH-----AGVLEGLSSNGCYLRPAAP-- 71
      |||:|||||
Db      111 TDSVVSQSAKKGSKRVHSNQHPOS PAVPPTYPGSGPPAASALSTPGGNGVPAAPAPS 170
      |||:|||||

QY      72 --GGIANPEKKNCGTCOCNSONTLS----SGPLT 99
      |||:|||||
Db      171 ALGPKASPAKSHNSGTBPAPYQAQAVAPPAAPSGPST 204
      |||:|||||

```

RESULT 5

bifocal protein- fruit fly (Drosophila melanogaster)
 C/Species: Drosophila melanogaster
 C/Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #ext_change 09-Jul-2004
 C/Accession: T03743
 R/BahrI, S.M.; Yang, X.Y.; Chia, W.
 Mol. Cell. Biol. 17, 5521-5529, 1997
 A/Title: The Drosophila bifocal gene encodes a novel protein which colocalizes with actin
 A/Reference number: Z15048; MUID:97415628; PMID:9271427
 A/Accession: T03743
 A/Status: preliminary; translated from GB/EMBL/DDbU
 A/Molecule type: mRNA
 A/Residues: 1-1063 <BAH>
 A/Cross-references: UNIPROT:O16125; EMBL:AF011791; NID:g2388667; PIDN:AA869991.1; PID:g2388667
 C/Genetics:
 A/Cross-references: Flybase:FBgn0014133
 A/Note: bifocal

Query Match

	Matches	39; Conservative	24; Mismatches	58; Indels	67; Gaps	6
Qy	5	GSRADAIERYRYESWTR-----	-ETSEFWLTYTSDAL-----	PSAAT-	42	
		: : :	: : :			
Db	83	GAIDFTEPATISSTSQKRMIGSEBEKEKSISINTNSDSTGCHSVAAVLSPPDAATT	142			
Qy	43	-----	DSGEAGGLHAGVLDELGSNGVLRLPAPG	72		
		: : :	: : :			
Db	143	NNTVTPIPKÖRSSLLNTRSQREREMRYTLSSGRDDELSGEGRPAGIVSNSKGVEYG	202			
Qy	73	GIANPEKKNNCGTCQPNSQNLSSGPLTOKÖNGILWATEAKDDAYKMSAREVAIVNTENIRO	132			
		: : :	: : :			
Db	203	TIGSPSSAN--QNPNPHLK----TKCKGGGVAGGKPSAKE-----	TIVDNKS	247		
Qy	133	MDSRKRVY	140			
		: : : :				
Db	248	CSKTYSIS	255			

RESULT 6

regulatory protein area - *Emricicella nidulans*
C/Species: *Emricicella nidulans*, *Aspergillus nidulans*
C/Date: 26-Jul-1996 #sequence revision 26-Jul-1996 #text_change 09-Jul-2004
C/Accession: A57988; S10017; S70168; S72883

R;Kudla, B.; Caddick, M.X.; Langdon, T.; Martinez-Rossi, N.M.; Bennett, C.F.; Sibley, S.
EMBO J. 9, 1355-1364, 1990
A>Title: The regulatory gene *areA* mediating nitrogen metabolite repression in *Aspergillus*
A/Reference number: S10017; MUID:90228331; PMID:1970293
A/Accession: A57988
A/Status: translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-876 <KUD>
A/Cross-references: UNIPROT:P17429; EMBL:X52491; NID:g1019911; PIDN:CAA36731.1; PID:g115
A/Note: this sequence represents reinterpretation to include two exons
A/Accession: S10017
A/Molecule type: DNA
A/Residues: 158-876 <KUD>
A/Cross-references: EMBL:X52491
A/Note: this sequence represents the authors' original translation
R;Langdon, T.; Sheerin, A.; Ravagnani, A.; Gielkens, M.; Caddick, M.X.; Arst Jr., H.N.
Mol. Microbiol. 17, 877-888, 1995
A>Title: Mutational analysis reveals dispensability of the N-terminal region of the *Aspe*
A/Reference number: S70167; MUID:96123430; PMID:8596437
A/Accession: S70168
A/Status: nucleic acid sequence not shown
A/Molecule type: DNA
A/Residues: 1-791, 'A', '793', 'T', '795', 'SPGTNS', '802-876 <LAN>
A/Cross-references: EMBL:X52491
R;Caddick, M.X.
submitted to the EMBL Data Library, October 1995
A/Reference number: S72883
A/Accession: S72883
A/Molecule type: DNA
A/Residues: 1-876 <CAD>
A/Cross-references: EMBL:X52491; NID:g1019911; PIDN:CAA36731.1; PID:g1154625
C/Genetics:
A/Gene: *areA*
A/Intons: 147/2
C/Function:
C/Description: mediates nitrogen metabolite repression
C/Superfamily: nitrogen regulatory protein ntr-2; GATA-type zinc finger homology
C/Keywords: DNA binding; transcription regulation; zinc finger
F;673-723/Domain: GATA-type zinc finger homology <GZF>
F;673-697/Region: zinc finger GATA motif

Query Match 10.3%; Score 79; DB 1; Length 876;
Best Local Similarity 28.4%; Pred. No. 22;
Matches 31; Conservative 10; Mismatches 28; Indels 40; Gaps 5;

QY 57 EDGLSSNGVLRLPAPGSIANPEKK-----MNCGTQC-----PNSQNLSSG----- 96
DB 645 ESGINSAPRPARPSPGCKNGEQNGPTTCNCFQTTPLWRNDEGQPLCNACGLFLKHL 704
QY 97 -----PLTKONGLMATEAKDKRMSAREVALINVTENIRQMDSKRYTK 141
DB 705 GVVRRPLSLK-----TDVIRKRNNSANSLAVG-----SSRSVK 737

RESULT 7
conserved hypothetical protein yobL - *Bacillus subtilis*
C/Species: *Bacillus subtilis*
C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C/Accession: C69899
R;Kunert, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Bertet
C.; Bron, S.; Brouillet, S.; Brunech, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A/Authors: Foulger, C.D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Gallizi, A.; Gall
lech, J.; Harwood, C.R.; Hentau, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
Koeter, P.; Koningerstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidis, A.; Lardinois
A.; Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maue
Y., M.; Ogawa, K.; Ogata, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portelle
Ritter, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scallion
A.; Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Ser
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpestre, P.; Tognoni, A.; Tosato, V.; Uchiyama
T.; Wintere, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K.

A/Authors: Yoshikawa, H.F.; Zumestein, E.; Yoshikawa, H.; Danchin, A.
A>Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*
A/Reference number: A69580; MUID:98044033; PMID:9384377
A/Accession: C69899
A/Status: preliminary; nucleic acid sequence not shown; translation not shown
A/Molecule type: DNA
A/Residues: 1-600 <KUN>
A/Cross-references: UNIPROT:Q34330; GB:Z99114; GB:AL009126; NID:g2634230; PIDN:CAB13792
A/Experimental source: strain 168
C/Genetics:
A/Gene: *yobL*

Query Match 10.2%; Score 78.5; DB 2; Length 600;
Best Local Similarity 23.7%; Pred. No. 16;
Matches 40; Conservative 22; Mismatches 56; Indels 51; Gaps 7;

QY 9 DAIEPRYSESMTRE-----TESTWLYTSDALPSAATDSGPBAGLHAGVLDGLS 61
DB 342 DAISAAIEESYQDMVNGDAVSRSMVY---AIGSVAVAVGTG---AGAINKADA 393
QY 62 SNGVLRPAPAGIA-----NPEKKNCGTQC-----NSQNLSSGPT----- 99
DB 394 AGKVINKASQAGKKIKVDKIPDLLPYNPKYKALADNPYNNVDSQNLKNELLTNAKIP 453
QY 100 -----QKONGLMATEAKDKRMSAREVALINVTENIRQMDSKRY 139
DB 454 DGTREKPTGQKSPPLNKEKD---AYEIGKVKAKGVADVSRV 497

RESULT 8
T37073
hypothetical protein SCJ30.06c - *Streptomyces coelicolor*
C/Species: *Streptomyces coelicolor*
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T37073
R;Sanders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.
submitted to the EMBL Data Library, August 1999
A/Reference number: Z21621
A/Accession: T37073
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-269 <SAN>
A/Cross-references: UNIPROT:Q9S1V3; EMBL:AL109973; PIDN:CAB53301.1; GSPDB:GN00070; SCOB
A/Experimental source: strain A312)
C/Genetics:
A/Gene: SCOEDB:SCJ30.06c

Query Match 10.2%; Score 78; DB 2; Length 269;
Best Local Similarity 25.8%; Pred. No. 6.9;
Matches 25; Conservative 10; Mismatches 38; Indels 24; Gaps 4;

QY 11 IEBRYYESWT-----RETESTWLYTSDALPSAATDSGPBAGLHAGVLDGL 60
DB 176 LPEITHVEQWVYVHFCHRRPRRYLRSGGLTFSPFTVLPNRGCCSGSRTG----- 225
QY 61 SSNGVLRPAPAGSIANPEKKMNC--TQCPNSQNLSS 95
DB 226 GGGVVTGAGGIV-PGEACRCRAAHPYNAKMS 260

RESULT 9
G90128
elongation factor EF-2 [imported] - *Gulliardia theta nucleomorph*
C/Species: *nucleomorph Gulliardia theta*
A/Note: a nucleomorph is the vestigial nucleus of a eukaryotic endosymbiont
C/Date: 10-May-2001 #sequence_revision 10-May-2001 #text_change 09-Jul-2004
C/Accession: G90128
R;Douglas, S.; Zanner, S.; Fraunholz, M.; Beaton, M.; Penny, S.; Deng, L.T.; Wu, X.; Rei
Nature 410, 1091-1096, 2001
A>Title: The highly reduced genome of an enslaved algal nucleus.
A/Reference number: A99082; MUID:11323671; PMID:11323671
A/Accession: G90128
A/Status: preliminary

A:Molecule type: DNA
A:Residues: 1-848 <DOU>
A:Cross-references: UNIPROT:Q98S60; GB:AF083031; NID:gl3794345; PIDN:AAK39722.1; GSPDB:G
C:Genetics:
A:Gene: EF2
A:Map position: 3
A:Genome: nucleomorph
C:Superfamily: translation elongation factor 2, translation elongation factor Tu homolog
:Keywords: nucleomorph

Query Match	10.1%	Score 77.5;	DB 2;	Length 848;
Best Local Similarity	25.2%;	Pred. No. 30;		
Matches	38;	Conservative	23;	Mismatches 65;
				Gaps 7;

```

OY 10 AIEIRYVESMTRETEBTWLTUTSDLPAAATDSGP-----AGLHAGV-----LEGLS 61
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 493 AVEKPNDSDPKULIEGL-KRLSKSDPLVQCMTEBSGEHVAAGAGELHETGLKDLQDDFM 551
OY 62 SNGVLRPAAP-----GGIANPEKKNMCGQCNSN-----LSSGLPTQCKNLMATEA 110
      |||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||::|||
Db 552 NGAELKISQIVSYRETVEGVNPESTAVCSKSPNKNRIYCAAEPLP-----GL--SEA 606

```

```
QY      111 KRDAKMSAREVAINTENTENRQMDRSKRVTK 141
        | | | : | : | : |
Db      607 IDDGKIKSSDEPKTRIKELQKEDMNEEDIK 637
```

RESULT 10
T09144

probable guanine nucleotide exchange factor RhogEF2 - fruit fly (*Drosophila melanogaster*)
N:Alternate names: Shar pei/DRhogEF2

C:species: Drosophila melanogaster
C:Date: 11-Jun-1999 #sequence 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: mo0144, mo0323

submitted to the EMBL Data Library, October 1997
R;Haecker, U.; Perrimon, N.
C/Accession: 105144; 105225

A;Reference number: Z16586
A;Accession: T09144

A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA

A:Residues: 1-2559 <HAE>
A:Cross-references: UNIPROT:O44113, EMBL:AF031930, NID:g2687355, PID:g2687356
R:Barrett, K.; Leptin, M.; Settlement, J.
Cell 91, 905-915, 1997

A; Title: The Rho GTPase and a putative RhoGEF mediate a signaling pathway for the cell
A; Reference number: Z16618; MUID:98088790; PMID:9428514

A:Accession: T09223
A:Status: preliminary; translated from GB/EMBL/DBJ

RESULT 11

hypothetical protein SPCC70.05c - fission yeast (*Schizosaccharomyces pombe*)
141551

C:Accession: TA1551
C:Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2000
C:Species: schizosaccharomyces pombe

R/wood, V.; Rajandream, M.A.; Barrell, B.G.; Seeger, K.; Harris, D. submitted to the EMBL Data Library, June 1998

A;Reference number: Z22001
A;Accession: T41551

A;Status: preliminary; translated from GB/EMBL/
A;Molecule type: DNA

A;Residues: 1-781 <W00>
A;Cross-references: UNIPROT:O74526; EMBL:AL023794; PDB:1CA1355.1; GSPDB:GN00066; SPDB
A;Experimental source: strain 972h-; cosmid c70
C;Genetics:

A;Gene: SPDB:SPCC70.05c
A;Map position: 3

Query Match	10.0%	Score 77;	DB 2;	Length 781;
Post-Local Simulation	26.6%	Score 77;	DB 2;	Length 781;

best local similarity 20.0%; freq. no. 50;
Matches 34; Conservative 16; Mismatches 64; Indels 14; Gaps 5

Qy 16 YESWIRETESTLWTYTDSDAL---PSAATTS-----GEAGGLHGVIEDGLSSNGV 65
||| | : ||| : : :
Db 326 YESWPSTEFDMFYAVSGSLKLTLPQGTFDCINPAMPSPFGYGKSKMSKDDNVGSSAN 385

QY 66 LRPAPGGIANPEKMNCG-TQCP--NSQNLSSGPLTQKQNGLMATEAKRDAKMSAREV 122

Db 386 TAPNSPTANSSEGNQNGPTTYPIKPPNTISEIPRLK-SGFIPYAKRVVRLSAYK 444

Qy	123	AINVENTI	130
		:: :::	
Db	445	LYDETKDM	452

RESULT 12

JE0301
inulinase (EC 3.2.1.7) - *Aspergillus niger*

CISpecies: Aspergillus niger
CDate: 05-Dec-1998 #sequence
CAccession: F00301
CRevision: 05-Dec-1998 #text_change
CDate: 09-Jul-2004

R. Ohta, K.; Akimoto, H.; Matsuda, S.; Toshimitsu, D.; Nakamura, T. *Cytocession*. 060501

A/Title: A family of Drosophila serotonin receptors with distinct intracellular signals
A/Reference number: S19155; MUID:92155185; PMID:1110937
A/Accession: S19156
A/Molecule type: mRNA
A/Residues: 1-645 <SAU>
A/Cross-references: UNIPROT:P28286; EMBL:Z11490; NID:g7506; PIDN:CAA77571.1; PID:g7507
C/Genetics:
A/Gene: FlyBase:5-HT1B
A/Cross-references: FlyBase:Fgn0004572
C/Superfamily: Octopamine receptor type I
C/Keywords: G protein-coupled receptor; glycoprotein; transmembrane protein
F/124-145/Domain: transmembrane #status predicted <TM1>
F/156-177/Domain: transmembrane #status predicted <TM2>
F/193-214/Domain: transmembrane #status predicted <TM3>
F/224-256/Domain: transmembrane #status predicted <TM4>
F/284-305/Domain: transmembrane #status predicted <TM5>
F/564-587/Domain: transmembrane #status predicted <TM6>
F/597-619/Domain: transmembrane #status predicted <TM7>

Query Match 9.8%; Score 76; DB 2; Length 645;
Best Local Similarity 31.0%; Pred. No. 30;
Matches 31; Conservative 12; Mismatches 31; Indels 26; Gaps 4;

QY 29 TYTDSALPSAATDSGPEAGLHAGVLEGLSSNGVLRPAAPG-----IANPEKKNCG 84
DB 489 TTTPEKALSGAGTAVAGVAGSGSGSGEGAGTEGNAGVGLGVLASIANPHOKL---- 545
QY 85 TCCPNSONLSSGPLTOKONGLMATEAKRDARMSAREVAI 124
DB 546 -----AKRRQL--LEAKRE--RKAQGLAI 566

RESULT 14

AH0784
Probable transport system periplasmic binding protein STY2452 [imported] - Salmonella en
C/Species: Salmonella enterica subsp. enterica serovar Typh
A/Note: this species has also been called Salmonella typhi
C/Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C/Accession: AH0784
R/Parikh, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Conerton, P.; Cronin, A.; Davies, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,
S.; Moule, S.; O'Garra, P.
Nature 413, 848-852, 2001
A/Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;
A/Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov
A/Reference number: AB0502; MUID:21534947; PMID:11677608
A/Accession: AH0784
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-601 <PAR>
A/Cross-references: GB:AL513382; PIDN:CAD02598.1; PID:G16503454; GSPDB:GN00176
C/Genetics:
A/Gene: STY2452
C/Superfamily: periplasmic oligopeptide-binding protein

Query Match 9.7%; Score 74.5; DB 2; Length 601;
Best Local Similarity 22.8%; Pred. No. 38;
Matches 33; Conservative 20; Mismatches 53; Indels 39; Gaps 7;

QY 15 YVESWTR-----ETESTWLTYTDSDA-----LPSAATD--SGPEAGLHAGVL 56
DB 342 FYNAMSRTSYFONTEYAANNYPADDELVLAPMKDLPEVFTQIYQPVVNSG----- 395
QY 57 EDGLSSNGVLRPA-----PGIANPEKKNCGTCCPNSONLSSGPLTOKONGLMATEAKR 112
DB 396 -DGYDRENILKADALLTQAGVAVINGQQRVNSVTKPLTFELL--LPASSNSQWVLPFOH 451
QY 113 DAKMSAREVAIVNTENIRQMDRSK 137
DB 452 NLOR-----LGITWTRQVDSNQ 469

RESULT 15

T14260
Period protein Per3 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 20-Sep-1999 #sequence_revision 20-Sep-1999 #text_change 09-Jul-2004
C/Accession: T14260
R/Zylka, M.J.; Shearman, L.P.; Weaver, D.R.; Reppert, S.M.
Neuron 20, 1103-1110, 1998
A/Title: Three period homologs in mammals: differential light responses in the suprach
A/Reference number: Z17943; MUID:98318231; PMID:9655499
A/Accession: T14260
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: mRNA
A/Residues: 1-1113 <ZYL>
A/Cross-references: UNIPROT:O70361; EMBL:AF050182; NID:g3136149; PID:g3136150; PIDN:AMC4
C/Genetics:
A/Gene: Per3
C/Keywords: circadian rhythm

Query Match 9.6%; Score 74; DB 2; Length 1113;
Best Local Similarity 25.0%; Pred. No. 87;
Matches 38; Conservative 22; Mismatches 70; Indels 22; Gaps 7;

QY 1 MCGGGS-----RADAIEPRYVESWTRTESTWLTYTDSALPSAATDSGPEAGL 51
DB 860 LGAAGSELAPLVPAAPPEPTTSGSQRVENNEAH--GBELPFISRSSP----L 913
QY 52 HAGVLEGLSSNGVLRPAAPPGIANPEKKNCG--GTCCPNSONLSSGPL--TOKONGLM 106
DB 914 QLNLLQEEHPAPSESADAVRGC-AGPDAGKHCVTGPSSGSRSHCTSGELATATAHQESAA 972
QY 107 ATEAKRDARMSAREVAIVNTEN-IRQMDRSK 137
DB 973 ASGSASSTYFSSTDYASEVSENRRQPPDROR 1004

Search completed: April 12, 2005, 15:44:46
Job time : 22.6667 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 12, 2005, 15:16:23 ; Search time 94.6667 Seconds

(without alignment)
784.347 Million cell updates/sec

Title: US-10-705-716a-8

Perfect score: 767
Sequence: 1 MCGGSRRAAIEPRYYESWT.....VTENIRQMDRSKRYTKNCIN 145

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03:*

1: uniprot_sprot:*

2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	767	100.0	145	2	Q8VHV1 mus musculu
2	746	97.3	145	2	Q920K5
3	649	84.6	145	2	Q9HA93
4	634	82.7	145	2	Q8WNE9
5	621.5	81.0	180	2	Q8WXS3
6	480	62.6	123	2	Q9CYS9
7	472.5	61.6	149	2	Q8WXS1
8	294	38.3	54	2	Q8VBS8
9	294	38.3	54	2	Q790N3
10	288	37.5	73	2	Q8WXS0
11	278	36.2	54	2	Q8WTP6
12	272	35.5	80	2	Q8WXS2
13	267	34.8	54	2	Q8WNE8
14	161	21.0	32	2	Q801V5
15	87.5	11.4	1073	1	PVDA_PLAKN
16	86	11.2	1449	2	Q7XPB1
17	85	11.1	882	1	AREA_ASPNG
18	84.5	11.0	367	2	Q7K2P9
19	84.5	11.0	494	2	Q8IR46
20	84.5	11.0	495	2	Q961T1
21	84.5	11.0	591	2	Q83H42
22	84.5	11.0	591	2	Q83NY7
23	84.5	11.0	659	2	Q9VY16
24	84.5	11.0	773	2	Q9FNV4
25	84	11.0	253	2	Q661S5
26	84	11.0	365	1	POLG_SUNVS
27	84	11.0	1433	1	CAT8_YEAST
28	83.5	10.9	291	2	Q871V5
29	83.5	10.9	865	1	NRFA_PENUR
30	83	10.8	476	2	Q7PV61
31	83	10.8	585	2	Q7UZ29

32	83	10.8	860	1	AREA_PENRO	013508 penticillium
33	81.5	10.6	214	2	Q8WTC2	Q8mc2 leucophaea
34	81.5	10.6	1100	2	Q91IA3	Q91ia3 white spot
35	81	10.6	219	2	Q82AV5	Q82av5 streptomyc
36	81	10.6	228	2	Q91TQ3	Q91tq3 tupalid her
37	81	10.6	516	2	Q74641	Q74641 aspergillus
38	81	10.6	1070	1	PVDA_PLAKN	P50494 plasmodium
39	80.5	10.5	155	2	Q6U5F5	Q6u5f5 plasmodium
40	80.5	10.5	852	1	AOR2_HUMAN	Q60341 homo sapien
41	80.5	10.5	2516	1	CCAD_DROME	Q24270 drosophila
42	80.5	10.5	2519	2	Q81P25	Q81p25 drosophila
43	80.5	10.5	2552	2	Q81P24	Q81p24 drosophila
44	80	10.4	124	2	Q868X0	Q868x0 plasmodium
45	80	10.4	516	1	INUZ_ASPFI	Q94220 aspergillus

ALIGNMENTS

RESULT 1

ID	Q8VHV1	PRELIMINARY	PRT	145 AA.
AC	Q8VHV1			
DT	01-MAR-2002 (TREMblrel. 20, Created)			
DT	01-MAR-2002 (TREMblrel. 20, Last sequence update)			
DT	25-OCT-2004 (TREMblrel. 28, Last annotation update)			
DE	BALC isoform 1-6-8 (Mus musculus 16 days neonate cerebellum cDNA, RIKEN full-length enriched library, clone:9630028H6 product:drain and DE acute leukemia, cytoplasmic, full insert sequence).			
GN	Name=BALC;			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euleleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=12956/SVEVAC;			
RX	MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;			
RA	Tanner S.M., Austin J.L., Leone G., Rush J.J., Flagg C., Heinonen K., Mrozek K., Sill H., Knuutila S., Kolitz J.E., Archer K.J.,			
RA	Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;			
RT	BALC, the human member of a novel mammalian neuroectoderm gene lineage, is implicated in hematopoiesis and acute leukemia.";			
RT	Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).			
RL	[2]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RP	STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RX	MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;			
RA	RIKEN FANTOM Consortium;			
RT	"Functional annotation of a full-length mouse cDNA collection.";			
RT	Nature 409:685-690(2001).			
RL	[4]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RP	STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RA	The FANTOM Consortium;			
RT	"The RIKEN Genome Exploration Research Group Phase I & II Team;			
RT	"Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs.";			
RT	Nature 420:563-573(2002).			
RL	[5]			
RN	SEQUENCE FROM N.A.			
RC	STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RP	STRAIN=C57BL/6J; TISSUE=Cerebellum;			
RX	MEDLINE=2049374; PubMed=11042159; DOI=10.1101/97.145100;			
RA	Carninci P., Shibata Y., Hayata N., Sugahara Y., Shibata K., Itoh M., Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;			
RT	"Normalization and subtraction of cap-trapper-selected cDNAs to			

```

RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoaka S., Sasaki N., Carninci P.,
RA Kondo H., Akiyama J., Nishi K., Kitenai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishino T., Harada A.,
RA Yamamoto S., Matsunoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaka S., Inoue K., Togawa Y., Izawa M., Ohara E., Wataniki K.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuda S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RL sequencing pipeline with 384 multicapillary sequencer.";
RN Genome Res. 10:1757-1771(2000).
RP [7]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RA Aachai J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hiraoka T., Hirozane T.,
RA Hori F., Imocani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Katon H., Kawai J., Kojima Y., Kondo S., Kono H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohnato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shingawa A., Shiraki T., Sogabe Y., Teganai M.,
RA Tagawa A., Takahashi F., Takaku-Akahira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yasunishi A., Muramatsu M., Hayashizaki Y.;
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF371320; AAL50516.1; -.
DR EMBL; AK079337; BAC37611.1; -.
DR MGD; MGI:1928704; Baa1c.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 15515 MW; 4972670A618C4D6D CRC64;

Query Match 100.0%; Score 767; DB 2; Length 145;
Best Local Similarity 100.0%; Pred. No. 5, 1e-65;
Matches 145; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIEPRYYESWTRETESWTLYTSDALPSAAATDSGPBAGGLHAGVLEDDL 60
DB 1 MCGGSRADAIIEPRYYESWTRETESWTLYTSDALPSAAATDSGPBAGGLHAGVLEDDL 60
QY 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNSLSSGPLETOKONGIMATKADAKMSAR 120
DB 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNSLSSGPLETOKONGIMATKADAKMSAR 120
QY 121 EVAINTENIRQMDRSKRYTKNCIN 145
DB 121 EVAINTENIRQMDRSKRYTKNCIN 145

RESULT 2
Q920K5 PRELIMINARY; PRT; 145 AA.
ID O920K5;
AC O920K5;
DT 01-DEC-2001 (TrEMBLrel. 19, Created)
DT 05-JUL-2004 (TrEMBLrel. 19, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DR Dem-A20-4 (BAALC isoform 1-6-8).
GN Name=Dem-A20-4; Synonyms=Baalc;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA Wang X., Tian Q., Li W., Okano A., Suzuki T.;
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.

```

```

RC STRAIN=Sprague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas R.C., Heimonen K.,
RA Mrozek K., Sill H., Kruutilla S., Kolitz J.E., Archer C.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neurocytokermin gene
RL lineage, is implicated in hematopoiesis and acute leukemia.";
RN Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AB073319; BAB70507.1; -.
DR EMBL; AF371321; AAL50517.1; -.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 145 AA; 15475 MW; D5A27AD67456F341 CRC64;

Query Match 97.3%; Score 746; DB 2; Length 145;
Best Local Similarity 97.2%; Pred. No. 5, 1e-63;
Matches 141; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIIEPRYYESWTRETESWTLYTSDALPSAAATDSGPBAGGLHAGVLEDDL 60
DB 1 MCGGSRADAIIEPRYYESWTRETESWTLYTSDALPSAAATDSGPBAGGLHAGVLEDDL 60
QY 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNSLSSGPLETOKONGIMATKADAKMSAR 120
DB 61 SSSNGVLRPAAPGCIANPEKKNCGTCCPSQNSLSSGPLETOKONGIMATKADAKMSAR 120
QY 121 EVAINTENIRQMDRSKRYTKNCIN 145
DB 121 EVAINTENIRQMDRSKRYTKNCIN 145

RESULT 3
Q9HA93 PRELIMINARY; PRT; 145 AA.
ID Q9HA93;
AC Q9HA93;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DR Hypothetical protein FLJ12015 (BAALC isoform 1-6-8) (Brain and acute
DE leukemia, cytoplasmic) (BAALC 1-6-8).
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole embryo;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J., Saito K., Kawai Y., Isono Y., Nakamura Y., Nagahari K.,
RA Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M., Shiratori A.,
RA Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H., Shimizu K.,
RA Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E., Omura Y.,
RA Abe K., Kamihara K., Katsura N., Sato K., Tanikawa M., Yamazaki M.,
RA Ninomiya K., Ishibashi T., Yamashita H., Murakawa K., Fujimori K.,
RA Tanai H., Kimata M., Watanabe M., Hiraoa S., Chiba Y., Ishida S.,
RA Ono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T., Kusano J.,
RA Kanehori K., Takahashi F., Fujii A., Hara R., Tanase T., Nomura Y.,
RA Togiya S., Komai F., Hara R., Takeuchi K., Arita M., Imose N.,
RA Musashino K., Yuki H., Oshima A., Sasaki N., Aotaka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Woriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujitani T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Omoto Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,

```

RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuno Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Isegai T., Sugano S.;
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [2]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tamer S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
 RA Calligaris M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.B.,
 RA Brownstein M.J., Uadin T.B., Toshyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mallary S.J.,
 RA Bosak S.A., Mcwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting M., Helton E., Kettelman M., Madan A.C., Rodchenko V., Sanchez A.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W., Butterfield Y.S.,
 RA Krzywinski M.I., Skalski U., Smalios D.E., Scherch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK022077; BAB13960.1; -;
 DR EMBL; AF371319; AAL50515.1; -;
 DR EMBL; BC011517; AAL1517.1; -;
 DR EMBL; AF363578; AAL50377.1; -;
 DR InterPro; IPR009728; BALC_N;
 DR Pfam; PF06989; BALC_N; 1.
 SQ SEQUENCE 145 AA; 15551 MW; CF892BBE283D92E CRC64;

Query Match 84.6%; Score 649; DB 2; Length 145;
 Best Local Similarity 83.4%; Pred. No. 8.5e-54;
 Matches 121; Conservative 10; Mismatches 14; Indels 0; Gaps 0;

QY 1 MCCGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSCPEAGGLHAGVLEDEL 60
 DB 1 MCCGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSCPEAGGLHAGVLEDEL 60
 QY 61 SSGNGVRPAAPGGINPEKKNCGTCPPSONLSSGPLETKOKNGLWATEAKRDAXKMSAR 120
 DB 61 PSNGVPRSTAPGGINPEKKNCGTCPPSONLSSGPLETKOKNGLWATEAKRDAXKMSAR 120
 QY 121 EVAIVNTENIRQMDRSKRYTKNCIN 145
 DB 121 EVTIVNTDSIQQDNRRTTKKCNV 145

RESULT 4
 Q8WNE9 PRELIMINARY; PRT; 145 AA.
 AC Q8WNE9;

DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
 DE BALC isoform 1-6-8.
 GN Name=BALC;
 OS Sus scrofa (Pig).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
 OX NCBI_TaxID=9823;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tamer S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
 RA Calligaris M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF371322; AAL50518.1; -;
 DR Pfam; PF06989; BALC_N; 1.
 SQ SEQUENCE 145 AA; 15401 MW; C75BD7D0E8F8E26 CRC64;

Query Match 82.7%; Score 634; DB 2; Length 145;
 Best Local Similarity 82.1%; Pred. No. 2.3e-52;
 Matches 119; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 1 MCCGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSCPEAGGLHAGVLEDEL 60
 DB 1 MCCGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSCPEAGGLHAGVLEDEL 60
 QY 61 SSGNGVRPAAPGGINPEKKNCGTCPPSONLSSGPLETKOKNGLWATEAKRDAXKMSAR 120
 DB 61 PSNGVPRSTAPGGINPEKKNCGTCPPSONLSSGPLETKOKNGLWATEAKRDAXKMSAR 120
 QY 121 EVAIVNTENIRQMDRSKRYTKNCIN 145
 DB 121 EVTIVNTDSIQQDNRRTTKKCNV 145

RESULT 5
 Q8WKS3 PRELIMINARY; PRT; 180 AA.
 AC Q8WKS3;
 DT 01-MAR-2002 (TREMblrel. 20, Created)
 DT 01-MAR-2002 (TREMblrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMblrel. 26, Last annotation update)
 DE BALC 1-5-6-8.
 GN Name=BALC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tamer S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
 RA Calligaris M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF363578; AAL50379.1; -;
 DR Pfam; PF06989; BALC_N; 1.
 SQ SEQUENCE 180 AA; 19224 MW; 380183E0F188F684 CRC64;

Query Match 81.0%; Score 621.5; DB 2; Length 180;
 Best Local Similarity 67.2%; Pred. No. 4.5e-51;
 Matches 121; Conservative 10; Mismatches 14; Indels 35; Gaps 1;

QY 1 MCCGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSCPEAGGLHAGVLEDEL 53
 DB 1 MCCGSRADAIIPRYESWTRETSWTLYTDSALPSAAATDSCPEAGGLHAGVLEDEL 53

QY 54 -----GVLBDGLSSNGVLRPAAPGGIANPEKKNCGT 85
 DB 61 KIKAPTVSDSEGLFASAKMPLAVFSGHGLDGLPSNGVPRSTAPGGIPNPEKKNCT 120
 QY 86 QCPNSQNTSSGPTQKONGLMATEAKRDPAKMSAREVAINTENIQMDRSKRVTKNCIN 145
 DB 121 QCPNPQSSSGPLTQKONGLQTEAKRDPAKMSAREVAINTENIQMDRSKRVTKNCIN 180

RESULT 6
 Q9CYS9 PRELIMINARY; PRT; 123 AA.
 AC Q9CYS9;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, last annotation update)
 DE Mus musculus 10, 11 days embryo whole body cDNA, RIKEN full-length
 DE enriched library, clone:2810457D07 product:brain and acute leukemia,
 DE cytoplasmic, full insert sequence.
 GN Name=Baalc;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NC NCBITaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carinci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning.";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=21085560; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN PANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection.";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA The PANTOM Consortium;
 RT "The RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carinci P., Shibata Y., Hayata N., Sugahara Y., Shibata K., Itoh M.,
 RA Kono H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
 RT "Normalization and subtraction of cap-trapper-selected cDNAs to
 RT prepare full-length cDNA libraries for rapid discovery of new genes.";
 RL Genome Res. 10:1617-1630(2000).
 RN [5]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RX MEDLINE=20330913; PubMed=11076861; DOI=10.1101/gr.152600;
 RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Saeki N., Carinci P.,
 RA Kono H., Akiyama J., Nishi K., Kitsuai T., Taishi H., Itoh M.,
 RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishimi T., Harada A.,
 RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
 RA Fujiwara S., Inoue K., Togawa Y., Izawa M., Ohara E., Wachihi M.,
 RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
 RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
 RT "RIKEN integrated sequence analysis (RISA) system-384-format
 RT sequencing pipeline with 384 multiplexed capillary sequencer.";
 RL Genome Res. 10:1757-1771(2000).
 RN [6]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Whole body;
 RA Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,

RA Arakawa T., Bono H., Carinci P., Fukuda S., Fukunishi Y., Furuo M.,
 RA Hanganaki T., Hara A., Hayatsu N., Hiramoto K., Hirooka T., Hori F.,
 RA Imomani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
 RA Kawai J., Kojima Y., Kono H., Kouda M., Koy S., Kurihara C.,
 RA Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
 RA Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
 RA Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
 RA Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
 RA Teijima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
 RA Muramatsu M., Hayashizaki Y.;
 RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AK013358; BAB28808.1; -.
 DR MGI; MGI:1928704; Baalc.
 SQ SEQUENCE 123 AA; 13391 MW; 33DFCB5F8CFAA0AB CRC64;

Query Match 62.6%; Score 480; DB 2; Length 123;
 Best Local Similarity 98.9%; Pred. No. 8,2e-38;
 Matches 92; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 53 AGVLBDGLSSNGVLRPAAPGGIANPEKKNCGTQCPNSQNTSSGPTQKONGLMATEAKR 112
 DB 31 SGVLBDGLSSNGVLRPAAPGGIANPEKKNCGTQCPNSQNTSSGPTQKONGLMATEAKR 90

QY 113 DAKRMSAREVAINTENIQMDRSKRVTKNCIN 145
 DB 91 DAKRMSAREVAINTENIQMDRSKRVTKNCIN 123

RESULT 7
 Q8WX51 PRELIMINARY; PRT; 149 AA.
 AC Q8WX51;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, last annotation update)
 DE BAALC 1-5-6-7-8.
 GN Name=BAALC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 NC NCBITaxID=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plass C., Heinonen K.,
 RA Mrozek K., Sill H., Knutti S., Kolitz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
 RT "BAALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF163578; AAL50381.1; -.
 DR Genew; HGNC:14333; BAALC.
 DR Pfam; PF06989; BAALC_N; 1.
 SQ SEQUENCE 149 AA; 15577 MW; DB26C19969B91F6E CRC64;

Query Match 61.6%; Score 472.5; DB 2; Length 149;
 Best Local Similarity 65.3%; Pred. No. 5,3e-37;
 Matches 94; Conservative 3; Mismatches 12; Indels 35; Gaps 1;

QY 1 MCGGSRDAIEPRYRESWTRETSWTWLTYSDDALPSAAATDSCPEAGGILHA----- 53
 DB 1 MCGGSRDAIEPRYRESWTRETSWTWLTYSDDALPSAAATDSCPEAGGILHVSLEAKS 60

QY 54 -----GVLBDGLSSNGVLRPAAPGGIANPEKKNCGT 85
 DB 61 KIKAPTVSDSEGLFASAKMPLAVFSGHGLDGLPSNGVPRSTAPGGIPNPEKKNCT 120

QY 86 QCPNSQNTSSGPTQKONGLMATEAKRDPAKMSAREVAINTENIQMDRSKRVTKNCIN 145
 DB 121 QCPNPQSSSGPLTQKONGLQTEAKRDPAKMSAREVAINTENIQMDRSKRVTKNCIN 180

RESULT 8

Q8VBS8
ID Q8VBS8 PRELIMINARY; PRT; 54 AA.
AC Q8VBS8;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAALC isoform 1-8.
GN Name=BAALC;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=12986/SEVETAC;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371324; AAL50520.1; -
DR MGD; MGI:1928704; Baalc.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BFBECDB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETSTWLTYSDDLPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIERYESWTRETSTWLTYSDDLPSAATDGPAGGLHAG 54

RESULT 9
Q790N3 PRELIMINARY; PRT; 54 AA.
AC Q790N3;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BAALC isoform 1-8.
GN Name=BAALC;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=Sprague-Dawley;
RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF371325; AAL50521.1; -
DR InterPro; IPR0069728; BAALC_N.
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 54 AA; 5667 MW; FEE8C0BFBECDB829 CRC64;

Query Match 38.3%; Score 294; DB 2; Length 54;
Best Local Similarity 100.0%; Pred. No. 1.5e-20;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETSTWLTYSDDLPSAATDGPAGGLHAG 54
DB 1 MCGGSRADAIERYESWTRETSTWLTYSDDLPSAATDGPAGGLHAG 54

RESULT 10
Q8WXS0
ID Q8WXS0 PRELIMINARY; PRT; 73 AA.
AC Q8WXS0;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE BAALC 1-4-5-6-8.
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
DR EMBL; AF363578; AAL50380.1; -
DR Pfam; PF06989; BAALC_N; 1.
SQ SEQUENCE 73 AA; 7671 MW; 98DBC2E6E524A CRC64;

Query Match 37.5%; Score 288; DB 2; Length 73;
Best Local Similarity 91.4%; Pred. No. 7.8e-20;
Matches 53; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 MCGGSRADAIERYESWTRETSTWLTYSDDLPSAATDGPAGGLHAG 58
DB 1 MCGGSRADAIERYESWTRETSTWLTYSDDLPSAATDGPAGGLHAG 58

RESULT 11
Q8WTP6 PRELIMINARY; PRT; 54 AA.
AC Q8WTP6;
DT 01-MAR-2002 (TrEMBLrel. 20, Created)
DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE BAALC isoform 1-8 (BAALC protein) (BAALC 1-8).
GN Name=BAALC;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plase C., Heinson K.,
RA Mrozek K., Sill H., Knuutila S., Koltz J.E., Archer K.J.,
RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.;
RT "BAALC, the human member of a novel mammalian neuroectoderm gene
lineage, is implicated in hematopoiesis and acute leukemia.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Akenauer R.D., Collins F.S., Wagner L., Scheffen C.M., Schuler G.D.,
RA Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.U., Ueda T.B., Toshiyuki S., Carninci P., Franke C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,

RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skaleja U., Smalusz D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences.";
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Brain;
 RA Strausberg R.;
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AF371323; AAL50519.1; -;
 DR EMBL; BC035038; AAL50503.1; -;
 DR EMBL; AF635378; AAL50378.1; -;
 DR Pfam; PF06989; BAALC_N; 1.
 SQ SEQUENCE 54 AA; 5663 MW; FEF084EABED9B829 CRC64;

Query Match 36.2%; Score 278; DB 2; Length 54;
 Best Local Similarity 94.4%; Pred. No. 4.9e-19;
 Matches 51; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 54
 DB 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 54

RESULT 12
 Q8WX52 PRELIMINARY; PRT; 80 AA.

AC Q8WX52;
 DT 01-JUN-2003 (TREMBlrel. 20, Created)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE BAALC 1-2;
 GN Name=BAALC;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxId=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BAALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF635378; AAL50382.1; -;
 DR Pfam; PF06989; BAALC_N; 1.
 SQ SEQUENCE 80 AA; 8573 MW; 70BE9B91C3245E2D CRC64;

Query Match 35.5%; Score 272; DB 2; Length 80;
 Best Local Similarity 94.3%; Pred. No. 2.9e-18;
 Matches 50; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 53
 DB 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 53

RESULT 13

Q8WNE8 PRELIMINARY; PRT; 54 AA.

AC Q8WNE8;
 DT 01-MAR-2002 (TREMBlrel. 20, Created)
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE BAALC isoform 1-8.
 GN Name=BAALC;
 OS Sus scrofa (Pig).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
 OX NCBI_TaxId=9823;
 RN [1]
 RP SEQUENCE FROM N.A.

RX MEDLINE=21574584; PubMed=11707601; DOI=10.1073/pnas.241525498;
 RA Tanner S.M., Austin J.L., Leone G., Rush L.J., Plas C., Heimonen K.,
 RA Mrozek K., Sill H., Knutti S., Koltz J.E., Archer K.J.,
 RA Caligiuri M.A., Bloomfield C.D., de la Chapelle A.,
 RT "BAALC, the human member of a novel mammalian neuroectoderm gene
 RT lineage, is implicated in hematopoiesis and acute leukemia.";
 RL Proc. Natl. Acad. Sci. U.S.A. 98:13901-13906(2001).
 DR EMBL; AF371326; AAL50522.1; -;
 DR Pfam; PF06989; BAALC_N; 1.
 SQ SEQUENCE 54 AA; 5681 MW; 6538C3DABED9B825 CRC64;

Query Match 34.8%; Score 267; DB 2; Length 54;
 Best Local Similarity 92.6%; Pred. No. 5.4e-18;
 Matches 50; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 54
 DB 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 54

RESULT 14
 Q801V5 PRELIMINARY; PRT; 32 AA.

AC Q801V5;
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
 DE SI:ZC215113.3 (Novel protein similar to human brain and acute
 DE leukemia, cytoplasmic (BAALC)) (Fragment).
 GN Name=SI:ZC215113.3;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxId=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Sehra H.;
 RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL807244; CAD87801.1; -;
 DR InterPro; IPR009728; BAALC_N.
 DR Pfam; PF06989; BAALC_N; 1.
 FT NON_TER 32
 FT 32
 SQ SEQUENCE 32 AA; 3663 MW; 916A5445D263B7F7 CRC64;

Query Match 21.0%; Score 161; DB 2; Length 32;
 Best Local Similarity 90.6%; Pred. No. 3.4e-08;
 Matches 29; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 32
 DB 1 MCGGSRADAIPEPRYESTWRTSTWLTYYTDSALPSAAATDSCPEAGGLHAG 32

RESULT 15

PVDA PLAKN STANDARD; PRT; 1073 AA.

AC P22545;
 DT 01-AUG-1991 (Rel. 19, Created)
 DT 01-OCT-1996 (Rel. 34, Last sequence update)
 DT 29-MAR-2004 (Rel. 43, Last annotation update)
 DE Duffy receptor, alpha form precursor (Erythrocyte binding protein).
 OS Plasmodium knowlesi.
 OC Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 OX NCBI_TaxId=5850;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=92357776; PubMed=1496004;

RA Adams J.H., Sim B.K., Dolan S.A., Fang X., Kaslow D.C., Miller L.H.;
 RT "A family of erythrocyte binding proteins of malaria parasites.";
 RL Proc. Natl. Acad. Sci. U.S.A. 89:7085-7089(1992).
 RN [2]
 RP SEQUENCE OF 296-1073 FROM N.A.
 RX MEDLINE=91004213; PubMed=2170017; DOI=10.1016/0092-8674(90)90295-P;
 RA Adams J.H., Hudson D.E., Toril M., Ward G.E., Wellens T.E., Aikawa M.,
 Miller L.H.;
 RT "The Duffy receptor family of Plasmodium knowlesi is located within
 the micronemes of invasive malaria merozoites.";
 RL Cell 63:141-153(1990).
 CC -1- FUNCTION: Binds to the human erythrocytes Duffy blood group
 determinant.
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; M90466; AAA29602.1; -;
 DR EMBL; M68517; AAA29590.1; -;
 DR EMBL; M68518; AAA29591.1; -;
 DR PIR; A35970; A35970.
 DR InterPro; IPR008602; Duffy_binding.
 DR Pfam; PF05424; Duffy_binding; 1.
 KW Glycoprotein; Malaria; Multigene family; Receptor; Signal;
 KW Transmembrane.
 FT SIGNAL 1 21 Potential.
 FT CHAIN 22 1073 Duffy receptor, alpha form.
 FT DOMAIN 22 1007 Extracellular (Potential).
 FT TRANSMEM 1008 1029 Potential.
 FT DOMAIN 1030 1073 Cytoplasmic (Potential).
 FT CARBOHYD 134 134 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 179 179 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 202 202 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 252 252 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 348 348 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 679 679 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 746 746 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 779 779 N-linked (GlcNAc...) (Potential).
 FT CARBOHYD 788 788 N-linked (GlcNAc...) (Potential).
 SQ SEQUENCE 1073 AA; 120683 MW; 3965FC9F46B71808 CRC64;

Query Match 11.4%; Score 87.5; DB 1; Length 1073;
 Best Local Similarity 31.8%; Pred. No. 24;
 Matches 28; Conservative 13; Mismatches 44; Indels 3; Gaps 3;

QY 31 TSDALPSAAATDSGEAGAGLAVLEGLSSNGV-LRPAAPGIANPEKKMNC-GTQCP 88
 Db 647 TVSSDVPVSGKSDSPSTASHALAGSEVHNGTDEPKEDKADPQKDIKVKQKDT 706

QY 89 NSGNLSS-GLUTOKONGLMATEAKRPAK 115
 Db 707 DRRSGSLGPHHTDRAITLGETHMEKDT 734

Search completed: April 12, 2005, 15:42:12
 Job time : 96.6667 secs

This Page Blank (uspto)